

STIC-Biotech/ChemLib

154239

mgj

From: Walicka, Malgorzata
Sent: Monday, May 23, 2005 5:45 PM
To: STIC-Biotech/ChemLib

Please search SEQ ID NO:1, 3, 4, 5, 6 in the application No. 10/622,893.

Thank you in advance.

Malgorzata A. Walicka, Ph.D.
Patent Examiner
Art Unit 1652, Recombinant Enzymes
USPTO, Remsen Building, Room 2C76
400 Dulany St.
Alexandria, VA 22313
Mail Room 2C70
Tel. (571) 272-0944, fax (571) 273-0944

RECEIVED
MAY 24 2005
(STIC)

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: 5/27/05
Date Completed: 5/27/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: 1 AA#: 4
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: atop
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:47:36 ; Search time 3.47974 Seconds
(without alignments)
1189.648 Million cell updates/sec

Title: US-10-622-893A-1

Perfect score: 63

Sequence: 1 MGGSGDDDLAL 12

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1462099 seqs, 344972447 residues

Total number of hits satisfying chosen parameters: 1462099

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	100.0	12	17	US-10-622-893A-1
2	63	100.0	472	17	US-10-622-893A-5
3	49	77.8	102	16	US-10-767-701-46722
4	49	77.8	149	16	US-10-767-701-46955
5	48	76.2	64	14	US-10-080-170-503
6	48	76.2	64	16	US-10-080-170-503
7	48	76.2	64	16	US-10-468-356-503
8	48	76.2	230	16	US-10-767-701-40998
9	47	74.6	173	16	US-10-437-963-200561
10	47	74.6	230	16	US-10-437-963-172039
11	46	73.0	103	14	US-10-289-454-386
12	46	73.0	625	16	US-10-437-963-152224
13	46	73.0	926	16	US-10-437-963-174460

```

14 44 69.8 75 16 US-10-437-963-133890
15 44 69.8 359 16 US-10-437-963-116167
16 44 69.8 652 16 US-10-437-963-198055
17 44 69.8 757 16 US-10-437-963-181509
18 43 68.3 143 16 US-10-767-701-32972
19 43 68.3 209 15 US-10-425-114-62751
20 43 68.3 312 16 US-10-437-963-182122
21 43 68.3 344 16 US-10-437-963-141563
22 43 68.3 361 16 US-10-437-963-146049
23 43 68.3 361 16 US-10-437-963-182133
24 43 68.3 477 15 US-10-425-114-63251
25 43 68.3 646 16 US-10-437-963-161859
26 42 66.7 15 15 US-10-323-069A-136
27 42 66.7 15 17 US-10-851-965-136
28 42 66.7 16 14 US-10-010-160-61
29 42 66.7 17 14 US-10-010-160-66
30 42 66.7 57 16 US-10-437-963-144342
31 42 66.7 120 16 US-10-767-701-48120
32 42 66.7 125 16 US-10-437-963-111374
33 42 66.7 157 15 US-10-259-194A-120
34 42 66.7 178 16 US-10-437-963-104580
35 42 66.7 223 15 US-10-424-599-245391
36 42 66.7 232 15 US-10-424-599-193534
37 42 66.7 247 16 US-10-437-963-129496
38 42 66.7 275 16 US-10-437-963-107388
39 42 66.7 341 16 US-10-437-963-107394
40 42 66.7 404 15 US-10-282-122A-63165
41 42 66.7 417 17 US-10-732-923-13830
42 66.7 421 16 US-10-437-963-126589
43 42 66.7 447 15 US-10-374-780A-1140
44 42 66.7 447 15 US-10-412-699B-1338
45 42 66.7 447 16 US-10-437-963-110675

```

ALIGNMENTS

```

RESULT 1
US-10-622-893A-1
; Sequence 1, Application US/10622893A
; Publication No. US20050014935A1
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong-Sheng
; APPLICANT: Datta, Abhijit
; APPLICANT: Wang, Yuping
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: DETERMINATION OF GLYCATED PROTEINS
; FILE REFERENCE: 46692001300
; CURRENT APPLICATION NUMBER: US/10/622,893A
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 40%-100% identity to leader sequence
US-10-622-893A-1

```

```

Query Match 100.0%; Score 63; DB 17; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 MGGSGDDDLAL 12
Db 1 MGGSGDDDLAL 12

```

```

RESULT 2
US-10-622-893A-5
; Sequence 5, Application US/10622893A
; Publication No. US20050014935A1

```

```

Sequence 133890,
Sequence 116167,
Sequence 198055,
Sequence 181509,
Sequence 32972, A
Sequence 62751, A
Sequence 182122,
Sequence 141563,
Sequence 146049,
Sequence 182133,
Sequence 63251, A
Sequence 161859,
Sequence 136, App
Sequence 61, Appl
Sequence 66, Appl
Sequence 48120, A
Sequence 111374,
Sequence 120, App
Sequence 104580,
Sequence 245391,
Sequence 129496,
Sequence 193534,
Sequence 107388,
Sequence 107394,
Sequence 63165, A
Sequence 13830, A
Sequence 126589,
Sequence 1140, Ap
Sequence 1338, Ap
Sequence 110675,

```

```
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong-Sheng
; APPLICANT: Datta, Abhijit
; APPLICANT: Wang, Yuping
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: DETERMINATION OF GLYCATED PROTEINS
; FILE REFERENCE: 466992001300
; CURRENT APPLICATION NUMBER: US/10/622,893A
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric protein
US-10-622-893A-5

Query Match          100.0%; Score 63; DB 17; Length 472;
Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGGSGDDDDLAL 12
Db 1 MGGSGDDDDLAL 12

RESULT 3
US-10-767-701-46722
; Sequence 46722, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 46722
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 7554236.pep
US-10-767-701-46722

Query Match          77.8%; Score 49; DB 16; Length 102;
Best Local Similarity 90.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGSGDDDDLAL 11
Db 74 GGGSGDDDDAA 83

RESULT 4
US-10-767-701-46955
; Sequence 46955, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
```

```
; SEQ ID NO 46955
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C49_289.pep
US-10-767-701-46955

Query Match          77.8%; Score 49; DB 16; Length 149;
Best Local Similarity 90.0%; Pred. No. 3.3;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGSGDDDDLAL 11
Db 91 GGGSGDDDDAA 100

RESULT 5
US-10-080-170-503
; Sequence 503, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 503
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-503

Query Match          76.2%; Score 48; DB 14; Length 64;
Best Local Similarity 80.0%; Pred. No. 2;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGSGDDDDLAL 11
Db 10 GGGSGDDDDLAL 19

RESULT 6
US-10-080-170-503
; Sequence 503, Application US/10080170
; Publication No. US20040121322A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 503
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-503

Query Match          76.2%; Score 48; DB 16; Length 64;
Best Local Similarity 80.0%; Pred. No. 2;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```



```
Qy      2 GSGGDDDDLA 11
Db      10 GGGGDDDDIA 19

RESULT 7
US-10-468-356-503
; Sequence 503, Application US/10468356
; Publication No. US20040197896A1
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENEOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 05394.0019
; CURRENT APPLICATION NUMBER: US/10/468,356
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: 10/080,170
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 655
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 503
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-468-356-503

Query Match      76.2%; Score 48; DB 16; Length 64;
Best Local Similarity 80.0%; Pred. No. 2;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      2 GSGGDDDDLA 11
Db      10 GGGGDDDDIA 19

RESULT 8
US-10-767-701-40998
; Sequence 40998, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 40998
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C15329_1.pep
US-10-767-701-40998

Query Match      76.2%; Score 48; DB 16; Length 230;
Best Local Similarity 66.7%; Pred. No. 7.4;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 MGGGDDDDLA 12
Db      15 LGGAGDDGLAI 26

RESULT 9
US-10-437-963-200561
; Sequence 200561, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 200561
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_96018C.1.pep
US-10-437-963-200561

Query Match      74.6%; Score 47; DB 16; Length 173;
Best Local Similarity 72.7%; Pred. No. 8;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      2 GSGGDDDDLA 12
Db      154 GGDGDDDDAAM 164

RESULT 10
US-10-437-963-172039
; Sequence 172039, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 172039
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_70212C.1.pep
US-10-437-963-172039

Query Match      74.6%; Score 47; DB 16; Length 230;
Best Local Similarity 80.0%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      2 GSGGDDDDLA 11
Db      93 GGDGDDDDLS 102

RESULT 11
US-10-289-454-386
; Sequence 386, Application US/10289454
; Publication No. US20030157479A1
```

; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin
; APPLICANT: Jennings, Gary
; APPLICANT: Sonderesger, Ivo
; TITLE OF INVENTION: Antigen Arrays for Treatments of Allergic Eosinophilic Diseases
; FILE REFERENCE: 1700.0360001
; CURRENT APPLICATION NUMBER: US/10/289,454
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: US 60/396,636
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: PCT/IB02/00166
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 10/050,902
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/331,045
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 386
; TYPE: PRT
; LENGTH: 103
; ORGANISM: Mouse His-eotaxin-C1
US-10-289-454-386

Query Match 73.0%; Score 46; DB 14; Length 103;
Best Local Similarity 90.0%; Pred. No. 6.8;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GSGDDDDLAL 12
|||:|||||
Db 10 GSGDDDDKAL 19

RESULT 12
US-10-437-963-152224
; Sequence 152224, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 152224
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_52297C.1.pep
US-10-437-963-152224

Query Match 73.0%; Score 46; DB 16; Length 625;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSGDDDD 9
|||:|||||
Db 144 GSGDDDD 151

RESULT 13
US-10-437-963-174460
; Sequence 174460, Application US/10437963
; Publication No. US20040123343A1

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 174460
; LENGTH: 926
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72399C.1.pep
US-10-437-963-174460

Query Match 73.0%; Score 46; DB 16; Length 926;
Best Local Similarity 88.9%; Pred. No. 66;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GSGDDDDL 10
|||:|||||
Db 666 GSGDDDDL 674

RESULT 14
US-10-437-963-133890
; Sequence 133890, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 133890
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_35717C.1.pep
US-10-437-963-133890

Query Match 69.8%; Score 44; DB 16; Length 75;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGSAGDDDRVEL 12
|||:|||||
Db 1 MGSAGDDDRVEL 12

RESULT 15
US-10-437-963-116167
; Sequence 116167, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 116167
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_19694C.1.pep
US-10-437-963-116167

Query Match 69.8%; Score 44; DB 16; Length 359;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MGGSGDDDDLAL 12
:|:|:|:|:|
Db 44 VGGGEDGDLAL 55

Search completed: May 27, 2005, 14:15:01
Job time : 4.47974 secs

THIS PAGE BLANK (USPTO)

3/21/2003

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:28:05 ; Search time 4.55437 Seconds
(without alignments)
1019.048 Million cell updates/sec

Title: US-10-622-893A-1
Perfect score: 63
Sequence: 1 MCGSGDDDLAL 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	76.2	64	5	ABU05852 M. tuberc
2	46	73.0	96	7	ADK17286 Virus-lik
3	43	68.3	130	5	ABJ05339 Human PTH
4	43	68.3	300	8	ADQ36935 Cell prol
5	43	68.3	300	8	ADQ15603 Rice stre
6	43	68.3	378	6	AAO19572 M sterili
7	42	66.7	15	7	ADH52757 pBAD-TOPO
8	42	66.7	17	5	ABG68924 Lawsonia
9	42	66.7	17	5	ABG68919 Lawsonia
10	42	66.7	157	8	ADI45189 Rice isop
11	42	66.7	402	4	ABG12558 Novel hum
12	42	66.7	404	6	ABU35241 Protein e
13	42	66.7	409	8	ADL05698 M. catarr
14	42	66.7	447	8	ADI42677 Plant tra
15	42	66.7	447	8	ADO02925 Thaleare
16	42	66.7	581	4	ABB59833 Drosophil
17	42	66.7	730	4	ABB61435 Drosophil
18	42	66.7	784	4	ABB62115 Drosophil
19	42	66.7	858	8	ADS20230 Human agg
20	42	66.7	1138	7	ABO77206 Pseudomon
21	42	66.7	1926	4	AAG84915 Shrimp wh
22	41.5	65.9	21	5	AAT75830 Wheat xyl
23	41	65.1	32	5	ABG80719 Human Igg
24	41	65.1	32	8	ADI40815 pCep-BK-F
25	41	65.1	47	5	ABG80720 Human IgG

26	41	65.1	47	8	ADI40817 pCep-SP-E
27	41	65.1	61	4	AAU22386 Human car
28	41	65.1	61	7	ADE46354 Human car
29	41	65.1	61	8	ADJ07772 Human car
30	41	65.1	67	3	AAB53618 Human col
31	41	65.1	103	5	ABG94342 Human res
32	41	65.1	103	5	ABG80654 Human res
33	41	65.1	107	5	ABG94322 Mouse res
34	41	65.1	107	5	ABG80634 Processed
35	41	65.1	124	5	ABG94340 Mouse mPr
36	41	65.1	124	5	ABG80652 Mouse trv
37	41	65.1	124	7	ADD24200 mPrPt-EK-
38	41	65.1	130	5	ABG80722 Mouse Res
39	41	65.1	130	8	ADI40821 Resistin-
40	41	65.1	185	8	ADR09188 Human pro
41	41	65.1	216	7	ADJ68201 Human hea
42	41	65.1	350	5	ABG94339 Mouse mPr
43	41	65.1	350	5	ABG80651 Mouse pri
44	41	65.1	350	7	ADD24199 mPrPt-EK-
45	41	65.1	357	7	ADJ70832 Human hea

ALIGNMENTS

RESULT 1
ABU05852
ID ABU05852 standard; protein; 64 AA.
XX
AC ABU05852;
XX
DT 08-APR-2003 (first entry)
XX
DE M. tuberculosis and M. leprae marker protein #503.
XX
KW Mycobacterioses; survival; virulence; protective antigen; vaccine;
KW Mycobacterial disease; tuberculosis; leprosy.
XX
OS Mycobacterium tuberculosis.
OS Mycobacterium leprae.
PN W0200274903-A2.
XX
PD 26-SEP-2002.
XX
PF 22-FEB-2002; 2002WO-IB001973.
XX
PR 22-FEB-2001; 2001US-0270123P.
XX (INSP) INST PASTEUR.
XX Cole S;
XX WFI; 2002-759885/82.
XX
PT Identifying and selecting genes for survival or virulence of mycobacteria
PT by a comparative genomic analysis of the sequences of Mycobacterium
PT tuberculosis and M. leprae.
PS Claim 17; Page 714-715; 874pp; English.
XX
CC This invention relates to a novel method for identifying essential genes
CC for survival or virulence of mycobacteria species. The method comprises
CC aligning the genomic sequence of a first mycobacterium species on a
CC genomic sequence of a second mycobacterium species and selecting a
CC polynucleotide sequence that is highly conserved in both genomes with no
CC counterparts in other bacterial genomic sequences and that corresponds to
CC an essential gene for the survival or virulence of mycobacterium species.
CC The method of the invention is useful for detecting M. tuberculosis or M.
CC leprae infection. The method reduces the number of potential new targets
CC and protective antigens for new drugs and vaccine compositions to treat
CC and prevent mycobacterial diseases, particularly tuberculosis and
CC leprosy. The present sequence represents a marker protein from

CC Mycobacterium tuberculosis and Mycobacterium leprae identified using the
 CC method of the invention
 XX
 SQ Sequence 64 AA;

Query Match 76.2%; Score 48; DB 5; Length 64;
 Best Local Similarity 80.0%; Pred. No. 3.6;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGGDDDLA 11
 |||||:
 Db 10 GGGGDDDDIA 19

RESULT 2
 ADK17286
 ID ADK17286 standard; peptide; 96 AA.
 XX AC ADK17286;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Virus-like particle repetitive antigen array peptide.
 XX
 KW antiallergic; antiasthmatic; cytostatic; vaccine; virus-like particle;
 KW interleukin; IL-5; IL-13; eotaxin; repetitive antigen array;
 KW allergic eosinophilic disease; asthma; Hodgkin's lymphoma.
 XX
 OS Unidentified.
 OS
 FN WO2003040164-A2.
 XX
 PD 15-MAY-2003.
 XX
 PF 07-NOV-2002; 2002WO-EP012455.
 XX
 PR 07-NOV-2001; 2001US-0331045P.
 PR 18-JAN-2002; 2002US-00050902.
 PR 21-JAN-2002; 2002WO-IB000166.
 PR 19-JUL-2002; 2002US-0356636P.
 XX
 PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
 XX
 PI Bachmann M, Jennings G, Sonderegger I;
 XX
 DR WPI; 2003-441516/41.
 XX
 PT Composition comprising an ordered and repetitive antigen or antigenic
 PT determinant array, useful as a medicament, or for manufacturing a
 PT medicament for treating allergic eosinophilic diseases, e.g. asthma, or
 PT Hodgkin's lymphoma.
 XX
 PS . Disclosure; SEQ ID NO 400; 245pp; English.
 XX
 CC The invention relates to a composition comprising a virus-like particle
 CC and at least one antigen, which is a protein or peptide of interleukin
 CC (IL)-5, IL-13 or eotaxin and is bound to the virus-like particle, or a
 CC core particle with at least one first attachment site and at least one
 CC antigen with at least one second attachment site, where the antigen is a
 CC protein or peptide of IL-5, IL-13 or eotaxin. The second attachment site
 CC is an attachment site naturally or not naturally occurring with the
 CC antigen or antigenic determinant. The second attachment site is capable
 CC of association to the first attachment site, and where the antigen or
 CC antigenic determinant and the core particle interact through the
 CC association to form an ordered and repetitive antigen array. The
 CC compositions are useful as medicaments, or for manufacturing a medicament
 CC or a vaccine for treating allergic eosinophilic diseases, e.g. asthma, or
 CC Hodgkin's lymphoma and related diseases. This sequence is used to
 CC generate the compound of the invention.
 XX
 SQ Sequence 96 AA;

Query Match 73.0%; Score 46; DB 7; Length 96;

```

XX DE Cell proliferation-related polypeptide #45.
XX KW cell proliferation related polypeptide; cell proliferation; senescence;
XX KM differentiation; stress response.
XX OS Oryza sativa.
XX PN WO2004061122-A2.
XX PD 22-JUL-2004.
XX PF 23-DEC-2003; 2003WO-US041200.
XX PR 26-DEC-2002; 2002US-043656SP.
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX PI Cooper B;
XX DR WPI; 2004-534388/51.
XX KW New nucleic acid molecule encoding a cell proliferation-related
PT polypeptide, useful for modulating cell proliferation, senescence,
PT differentiation, development, and stress response in plants, and for
PT producing enhanced food crops.
XX Claim 28; SEQ ID NO 94; 408pp; English.
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a cell proliferation-related polypeptide. The nucleic acid
CC molecule and the encoded polypeptide, and methods are useful for
CC modulating cell proliferation, senescence, differentiation, development,
CC and stress response in plants, and for producing enhanced food crops. The
CC present sequence represents a cell proliferation-related polypeptide. The
CC present sequence is published separately from the main body of the
CC specification as EPO data.
XX SQ Sequence 300 AA;

Query Match 68.3%; Score 43; DB 8; Length 300;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MGGSGDDDDIAL 12
Db :|||:|||
46 IGGGGEDDDAL 57

RESULT 5
ADQ15603
ID ADQ15603 standard; protein; 300 AA.
XX AC ADQ15603;
XX DT 07-OCT-2004 (first entry)
XX DE Rice stress-related protein #6.
XX KW rice; stress-related protein; plant maturation; plant development;
XX plant proliferation; plant senescence; plant disease-resistance;
XX plant stress response; transgenic plant; pest tolerance;
XX herbicide tolerance; biotic stress tolerance; abiotic stress tolerance;
XX improved nutritional value; increased yield; increased proliferation.
XX OS Oryza sativa.
XX PN WO2004061080-A2.
XX PD 22-JUL-2004.
XX PF 23-DEC-2003; 2003WO-US041098.
XX PS Claim 24; Page 109-111; 116pp; Japanese.

PR 26-DEC-2002; 2002US-0436564P.
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX PI Cooper B;
XX DR WPI; 2004-534374/51.
XX N-PSDB; ADQ15602.
XX New isolated nucleic acids and proteins, useful for producing transgenic
PT plants having improved properties, e.g. tolerance to pests, herbicides,
PT or biotic or abiotic stresses, improved nutritional value, or increased
PT yield or proliferation.
XX Claim 28; SEQ ID NO 12; 551pp; English.
XX The invention comprises the amino acid and coding sequences of rice
CC stress-related proteins. The DNA and protein sequences of the invention
CC are useful for regulating and controlling plant maturation and
CC development, including proliferation, senescence, disease-resistance, or
CC stress response. They are also useful for producing transgenic plants
CC having improved properties, e.g. tolerance to pests, herbicides, or
CC biotic or abiotic stresses, improved nutritional value, increased yield
CC or proliferation, or improved structure causing less loss from lodging or
CC shattering. The present amino acid sequence represents a rice stress-
CC related protein of the invention.
XX SQ Sequence 300 AA;

Query Match 68.3%; Score 43; DB 8; Length 300;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MGGSGDDDDIAL 12
Db :|||:|||
46 IGGGGEDDDAL 57

RESULT 6
AAO19572
ID AAO19572 standard; protein; 378 AA.
XX AC AAO19572;
XX DT 28-JAN-2003 (first entry)
XX DE M sterilia protein fragment #7.
XX KW Substance PF1022; biosynthesis; transgenic; chorismic acid; anthelmintic;
XX phenylalanine; p-aminophenylpyruvic acid; pharmaceutical;
XX veterinary drug.
XX OS Mycelia sterilia.
XX PN WO200277244-A1.
XX PD 03-OCT-2002.
XX PF 22-MAR-2002; 2002WO-JP002782.
XX PR 22-MAR-2001; 2001JP-00082227.
XX (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX Yanai K, Sumida N, Watanabe M, Moriya T, Murakami T;
XX WPI; 2003-018934/01.
XX Novel biosynthesis gene-transferred transformants for producing PF1022
PT substance derivatives by fermentation, as pharmaceuticals or veterinary
PT drugs with anthelmintic activity.
XX Claim 24; Page 109-111; 116pp; Japanese.

```


QY 1 MG-GSGDDDD-LAL 12
 DB 1 MGS GSGDDDDKLAL 14

RESULT 9
 ABG68919
 ID ABG68919 standard; peptide; 17 AA.
 XX
 AC ABG68919;
 XX
 DT 10-OCT-2002 (first entry)
 XX
 DE Lawsonia intracellularis immunogenic protein related peptide #1.
 XX
 KW Immunogen; fihB; flir; ntrC; glnH; motA; motB; tlyC; ytfM; ytfN; porcine;
 KW pig; avian; bird; porcine proliferative enteropathy; PPE;
 KW intestinal adenomatosis complex; porcine intestinal adenomatosis; PIA;
 KW necrotic enteritis; proliferative haemorrhagic enteropathy;
 KW regional ileitis; haemorrhagic bowel syndrome; vaccine; antibacterial;
 KW porcine proliferative enteritis; Campylobacter spp.-induced enteritis.
 XX
 OS Synthetic.
 XX
 PN WO200238594-A1.
 XX
 PD 16-MAY-2002.
 XX
 XX 09-NOV-2001; 2001WO-AU001462.
 XX
 PR 10-NOV-2000; 2000AU-00001381.
 PR 17-NOV-2000; 2000US-0249596P.
 XX
 XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
 PA (AUPO-) AUSTRALIAN PORK LTD.
 PA (PFIZ) PFIZER PROD INC.
 XX
 PI Rosey EL, King KW, Good RT, Strugnell RA;
 XX
 DR WPI; 2002-557448/59.
 XX
 XX New immunogenic polypeptide comprising epitope of Lawsonia spp.
 PT polypeptide such as fihB, flir, ntrC, glnH, motA, polypeptides, useful in
 PT vaccines for treatment of porcine proliferative enteropathy in pigs and
 PT birds.
 XX
 PS Example 2; Page 152; 155pp; English.
 XX
 CC The invention describes an isolated or recombinant immunogenic
 CC polypeptide (I) which comprises, mimics or cross-reacts with a B-cell or
 CC T-cell epitope of a Lawsonia spp. polypeptide such as fihB, flir, ntrC,
 CC glnH, motA, tlyC, ytfM or ytfN polypeptides. (I) is useful for
 CC identifying whether or not a porcine or avian animal has suffered from a
 CC past infection, or is currently infected, with Lawsonia spp. or a
 CC microorganism that is immunologically cross-reactive with Lawsonia spp.
 CC Antibodies are useful for diagnosing infection of a porcine or avian
 CC animal by Lawsonia spp. or a microorganism that is immunologically cross-
 CC reactive with Lawsonia spp. A nucleic acid encoding a Lawsonia spp. or
 CC immunogen is useful as probes or primers for detecting Lawsonia spp. or
 CC related microorganism in a biological sample derived from a porcine or
 CC avian animal subject. (I) is preferably useful for vaccinating porcine
 CC animals against intestinal diseases collectively known as porcine
 CC proliferative enteropathy (PPE), previously known as intestinal
 CC adenomatosis complex, porcine intestinal adenomatosis (PIA), necrotic
 CC enteritis, proliferative haemorrhagic enteropathy, regional ileitis,
 CC haemorrhagic bowel syndrome, porcine proliferative enteritis and
 CC Campylobacter spp.-induced enteritis. (I) is also useful in vaccines for
 CC the prophylaxis and treatment of PPE in birds. This sequence represents a
 CC Lawsonia intracellularis immunogenic protein related peptide
 XX
 SQ Sequence 17 AA;

Query Match 66.7%; Score 42; DB 5; Length 17;
 Best Local Similarity 85.7%; Pred. No. 7.9;
 Matches 12; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 MG-GSGDDDD-LAL 12
 DB 1 MGS GSGDDDDKLAL 14

RESULT 10
 ADI45189
 ID ADI45189 standard; protein; 157 AA.
 XX
 AC ADI45189;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Rice isoprenoid biosynthesis-associated protein #60.
 XX
 KW Rice, isoprenoid biosynthesis; plant; isopentenyl diphosphate; IPP;
 KW dimethylallyl alcohol; DMAPP; short-chain plastid prenyltransferase;
 KW gibberellin; carotenoid; abscisic acid; tocopherol; plastoquinone;
 KW phyloquinone; mevalonate pathway; phytosterol; brassinosteroid;
 KW ubiquinone; monoterpene; sesquiterpene; protein prenylation; chlorophyll;
 KW haeme; yield.
 XX
 OS Oryza sativa.
 XX
 PN US2004010815-A1.
 XX
 PD 15-JAN-2004.
 XX
 XX 26-SEP-2002; 2002US-00259194.
 XX
 PR 26-SEP-2001; 2001US-0325277P.
 PR 04-APR-2002; 2002US-0370620P.
 PR 04-APR-2002; 2002US-0370743P.
 XX
 XX (LANG/) LANGE B M.
 PA (GHAS/) GHASSEMIAN M.
 PA (BRIG/) BRIGGS S P.
 PA (COOP/) COOPER B.
 PA (GLAZ/) GLAZEBROOK J.
 PA (GOFF/) GOFF S A.
 PA (KATA/) KATAGIRI F.
 PA (KREP/) KREPS J.
 PA (MOUG/) MOUGHAMER T.
 PA (PROV/) PROVART N.
 PA (RICK/) RICHE D.
 PA (ZHUT/) ZHU T.
 XX
 XX Lange BM, Ghassenian M, Briggs SP, Cooper B, Glazebrook J;
 PI Goff SA, Katagiri F, Kreps J, Moughamer T, Provart N, Ricke D;
 PI Zhu T;
 XX
 DR WPI; 2004-090562/09.
 DR N-PSDB; ADI45188.
 XX
 XX New isolated polynucleotides and polypeptides associated with isoprenoid
 PT synthesis in plants, useful for producing transgenic plants, for targeted
 PT gene disruption, as well as markers or probes.
 XX
 PS Claim 4; SEQ ID NO 120; 117pp; English.
 XX
 CC The invention relates to a polynucleotide (or its complement, protein
 CC encoding fragment or reverse complement), comprising a nucleotide
 CC sequence encoding a polypeptide comprising an amino acid sequence
 CC involved in or associated with the biosynthesis of isoprenoids in a rice
 CC plant. Also included are an isolated polypeptide involved in or
 CC associated with the biosynthesis of isoprenoids in a plant, an expression
 CC cassette comprising the polynucleotide, a host cell comprising the
 CC expression cassette, and a transgenic plant comprising the expression
 CC cassette. The polypeptides and polynucleotides include those associated

CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 404 AA;

Query Match 66.7%; Score 42; DB 6; Length 404;
 Best Local Similarity 87.5%; Pred. No. 2.1e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GSGGDDDD 9
 || |||||
 Db 381 GGGGDDDD 388

RESULT 13
 ADL05698
 ID ADL05698 standard; protein; 409 AA.

AC ADL05698;

DT 06-MAY-2004 (first entry)

DE M. catarrhalis protein #1464.

EX Moraxella catarrhalis; infection.

XX Moraxella catarrhalis.

OS US6673910-B1.

PN 06-JAN-2004.

PD 04-APR-2000; 2000US-00540236.

PF 08-APR-1999; 99US-0128416P.

PR (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL;

PI WPI: 2004-178127/17.

DR N-PSDB; ADL03778.

XX New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
 PT preparing a composition for diagnosing, preventing or treating infection

PT caused by Moraxella catarrhalis.
 XX Disclosure; SEQ ID NO 384; 429pp; English.

CC The invention relates to an isolated nucleic acid encoding an Moraxella
 CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
 CC composition for diagnosing, preventing or treating infection caused by
 CC Moraxella catarrhalis. The present sequence represents the amino acid
 CC sequence of a M. catarrhalis protein.

XX Sequence 409 AA;

Query Match 66.7%; Score 42; DB 8; Length 409;
 Best Local Similarity 87.5%; Pred. No. 2.1e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GSGGDDDD 9
 || |||||
 Db 386 GGGGDDDD 393

RESULT 14
 ADI42677
 ID ADI42677 standard; protein; 447 AA.

XX ADI42677;

XX 22-APR-2004 (first entry)

DE Plant transcription factor #408.

XX transgenic; plant; enhanced tolerance to abiotic stress;
 KW glyphosate tolerance; hormone sensitivity; disease resistance;
 KW sugar sensing; flowering; flower structure; stem bifurcation;
 KW branching pattern; apical dominance; trichome; stem morphology;
 KW root growth; root hair; seed development; cell proliferation;
 KW cell differentiation; premature senescence; necrosis; plant size;
 KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
 KW plant anthocyanin; light response; shade avoidance; bioinformatic;
 KW transcription factor; gene; ds.

OS Oryza sativa.

XX US2004019927-A1.

XX 29-JAN-2004.

PF 25-FEB-2003; 2003US-00374780.

XX 18-APR-2001; 2001US-00837944.

XX (SHER/) SHERMAN B K.

PA (RIEC/) RIECHMANN J L.

PA (JIAN/) JIANG C.

PA (HEAR/) HEARD J E.

PA (HAAR/) HAAKE V.

PA (CREE/) CREELMAN R A.

PA (RATC/) RATCLIFFE O.

PA (ADAM/) ADAM L J.

PA (REUB/) REUBER T L.

PA (KEDD/) KEDDIE J.

PA (BROU/) BROUN P E.

PA (PILG/) PILGRIM M L.

PA (DUBE/) DUBELL A N.

PA (PINE/) PINEDA O.

PA (YUGG/) YU G.

XX Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
 PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;
 PI Pilgrim ML, Dubell AN, Pineda O, Yu G;

XX WPI: 2004-132245/13.

PT New transgenic plant comprising a recombinant polynucleotide of any one
 PT of more than 500 nucleotide sequences, useful in bioinformatic search
 PT methods.

XX Claim 1; SEQ ID NO 1140; 435pp; English.

CC The invention describes a transgenic plant comprising a recombinant
 CC polynucleotide of any one of more than 500 nucleotide sequences fully
 CC defined in the specification or its complement. The method of the
 CC invention can be used to produced a plant having altered traits such as:
 CC enhanced tolerance to abiotic stress; glyphosphate tolerance; hormone
 CC sensitivity; disease resistance; sugar sensing; early or late flowering;
 CC altered flower structure, change in stem bifurcations, altered branching
 CC pattern, reduced apical dominance, reduced trichome density; lack of
 CC trichomes; reduced ectopic trichome development; altered trichome
 CC development; increase in trichome number; altered stem morphology;
 CC increased root growth; increased root hairs; altered seed development;
 CC altered cell proliferation or cell differentiation; rapid development;
 CC premature senescence; increased necrosis; increase in seedling or plant
 CC size; decreased plant size; leaf morphology; seed morphology; seed
 CC biochemistry; increase in root anthocyanins; increase in plant
 CC anthocyanins, or alteration in light response or shade avoidance. The
 CC transgenic plant, polynucleotides and polypeptides are useful in
 CC bioinformatic search methods. This is the amino acid sequence of a plant
 CC transcription factor, and an orthologue of Arabidopsis thaliana
 CC transcription factors isolated in the invention, that can be used in the
 CC creation of a transgenic plant with altered traits.

XX Sequence 447 AA;

Query Match 66.7%; Score 42; DB 8; Length 447;

Best Local Similarity 72.7%; Pred. No. 2.4e+02;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGSGDDDLAL 12

DB 75 GGGGDDDLGL 85

RESULT 15

AD002925
 ID ADO02925 standard; protein; 447 AA.

XX ADO02925;

XX 01-JUL-2004 (first entry)

DE Thalecress transcription factor, Rice orthologue #66.

XX Rice; transcription factor; plant; transgenic; abiotic stress;
 KW cold tolerance; heat tolerance; drought; osmotic stress;
 KW phosphate limitation; potassium limitation; nitrogen limitation;
 KW hormone sensitivity; disease resistance; sugar sensing; seed germination;
 KW flowering; inflorescence architectural change;
 KW meristem cell differentiation; phyllotaxy; apical dominance;
 KW trichome development; seed development; premature senescence;
 KW delayed senescence; lethality; necrosis; plant size; leaf morphology;
 KW seed morphology; secondary metabolism; light response; shade avoidance.

OS Oryza sativa.

XX US2004045049-A1.

PN 04-MAR-2004.

XX 10-APR-2003; 2003US-00412699.

XX 13-SEP-1999; 99US-00394519.

PR 21-JAN-2000; 2000US-00489176.

PR 17-FEB-2000; 2000US-00506720.

PR 22-MAR-2000; 2000US-00532591.

PR 22-MAR-2000; 2000US-00533029.

PR 22-MAR-2000; 2000US-00533030.

PR 22-MAR-2000; 2000US-00533392.
 PR 22-MAR-2000; 2000US-00533648.
 PR 06-APR-2000; 2000WO-US009448.
 PR 16-NOV-2000; 2000US-00713994.
 PR 27-MAR-2001; 2001US-00819142.
 PR 17-APR-2001; 2001US-00837444.
 PR 30-JAN-2002; 2002US-00958131.
 PR 14-JUN-2002; 2002US-00171468.
 PR 09-AUG-2002; 2002US-00225066.
 PR 09-AUG-2002; 2002US-00225067.
 PR 09-AUG-2002; 2002US-00225068.
 PR 17-DEC-2002; 2002US-0434166P.
 PR 25-FEB-2003; 2003US-00374780.

XX (ZHAN/) ZHANG J.

PA (FROM/) FROMM M E.

PA (HEAR/) HEARD J E.

PA (RIEC/) RIECHMANN J L.

PA (ADAM/) ADAM L J.

PA (BROU/) BROUN P E.

PA (PINE/) PINEDA O.

PA (REUB/) REUBER T L.

PA (KEDD/) KEDDIE J S.

PA (YUGG/) YU G.

PA (JIAN/) JIANG C.

PA (SAMA/) SAMAHA R S.

PA (PILG/) PILGRIM M L.

PA (CREE/) CREELMAN R A.

PA (DUBE/) DUBELL A N.

PA (RATC/) RATCLIFFE O.

PA (KUMI/) KUMIMOTO R.

PA (SHER/) SHERMAN B K.

XX Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE,

PI Pineda O, Reuber TL, Keddle JS, Yu G, Jiang C, Samaha RS,

PI Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;

PI Sherman BK;

XX WPI; 2004-225755/21.

PT New transgenic plant, useful in developing phenotypes with altered or
 PT improved characteristics or traits.

XX Claim 1; SEQ ID NO 1339; 213pp; English.

XX The invention relates to a transgenic plant comprises a recombinant
 CC polynucleotide having a polynucleotide sequence or its complementary
 CC sequence comprising a sequence encoding a polypeptide, that initiates
 CC transcription (i.e. a transcription factor) from Arabidopsis, Soybean,
 CC Rice, Rape or Corn, comprising any of the sequences appearing as ADO01588
 CC -ADO03527 or ADO03530-ADO03559. Also included are using a transgenic
 CC plant to grow a progeny plant, an expression cassette (comprising a
 CC constitutive, inducible or tissue-specific promoter and a recombinant
 CC polynucleotide described above), a host cell comprising the expression
 CC cassette, producing a modified plant having a modified trait, identifying
 CC a factor that is modulated by or interacts with a polypeptide encoded by
 CC the polynucleotide sequence and identifying at least one downstream
 CC polynucleotide sequence that is subject to a regulatory effect of any of
 CC the polypeptides encoded by the polynucleotide described above. The
 CC transgenic plant is useful for producing a plant that has an altered
 CC trait e.g. an enhanced tolerance to abiotic stress (increased tolerance
 CC to chilling, germination in cold conditions, freezing tolerance, tolerance
 CC to heat, tolerance to drought, tolerance to osmotic stress, tolerance to
 CC salt, tolerance to phosphate limitation, tolerance to potassium
 CC limitation, decreased sensitivity to nitrogen limitation), altered
 CC hormone sensitivity, reduced sensitivity to abscisic acid, an altered
 CC response to ethylene, disease resistance, altered susceptibility to
 CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to
 CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered
 CC susceptibility to Sclerotinia, altered sugar sensing, improved seed
 CC germination and seedling vigor, early flowering, late flowering, extended
 CC period of flowering, an inflorescence architectural change, a change in
 CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell

CC differentiation, altered phyllotaxy, altered branching pattern, reduced
 CC apical dominance, reduced trichome density, ectopic trichome development,
 CC altered trichome development, altered stem morphology, increased root
 CC growth, increased root hairs, altered seed morphology, altered cell
 CC proliferation/cell differentiation, premature senescence, delayed
 CC senescence, lethality, increased necrosis, an increase in seedling or
 CC plant size, decreased plant size, a change in leaf morphology, increased
 CC altered leaf development, increased leaf size and mass, glossy leaves,
 CC leaf cell expansion, change in seed morphology, altered seed coloration,
 CC increased seed size, decreased seed size, altered seed shape, change in
 CC leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid
 CC content, increased leaf insoluble sugars, decreased leaf insoluble
 CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid
 CC content, an alteration of leaf glucosinolate content, change in seed
 CC biochemistry, an increase in seed oil content, decrease in seed oil
 CC content, increase in seed fatty acid content, decrease in seed fatty acid
 CC content, increase in seed protein content, decrease in seed protein
 CC content, alteration in seed prenyl lipid content, increase in seed
 CC sterols, upregulation of genes involved in secondary metabolism, increase
 CC in root anthocyanins, increase in plant anthocyanins, and alteration in
 CC light response or shade avoidance. The present sequence represents an
 CC orthologue of a thalecress transcription factor isolated from Rice.
 XX
 SQ Sequence 447 AA;

Query Match 66.7%; Score 42; DB 8; Length 447;
 Best Local Similarity 72.7%; Pred. No. 2.4e+02;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGGGDDDDLAL 12
 |||||
 Db 75 GGGGDDDLAL 85

Search completed: May 27, 2005, 13:43:51
 Job time : 7.55437 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:34:21 ; Search time 0.921109 Seconds
(without alignments)
1253.491 Million cell updates/sec

Title: US-10-622-893A-1
Perfect score: 63
Sequence: 1 MGSGDDDDLA 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	76.2	64	2 B70512	hypothetical prote
2	44	69.8	240	2 H84708	hypothetical prote
3	43	68.3	894	2 T27007	hypothetical prote
4	42	66.7	585	2 S46825	hypothetical prote
5	42	66.7	730	2 A53064	folded gastrulatio
6	42	66.7	818	2 T02823	probable membrane
7	42	66.7	918	2 S44769	C29E4.2 protein -
8	42	66.7	943	2 S54493	hypothetical prote
9	42	66.7	1360	2 T18403	asparagine/asparta
10	41	65.1	84	2 B87376	hypothetical prote
11	41	65.1	489	2 T36100	probable ATP-bindi
12	41	65.1	508	1 ISH000	protein disulfide-
13	41	65.1	510	1 ISB050	protein disulfide-
14	41	65.1	815	2 T46169	CDC48-like protein
15	41	65.1	875	1 S66672	phosphatidylinoait
16	40	63.5	251	1 RNBV3C	DNA-directed RNA p
17	40	63.5	294	2 C82643	daunorubicin C-13
18	40	63.5	334	2 A29561	prostatic spermine
19	40	63.5	353	2 E95914	probable secreted
20	40	63.5	504	2 G84251	hypothetical prote
21	40	63.5	539	2 G95405	hypothetical prote
22	40	63.5	825	2 JS0174	cellulase [EC 3.2.
23	40	63.5	1074	2 T52654	Ca2+-transporting
24	40	63.5	1150	2 D88556	protein B0464.2 [i
25	40	63.5	1274	2 S28279	hypothetical prote
26	40	63.5	6420	2 T30283	polyketide synthas
27	39	61.9	58	2 S03991	Band 3 anion trans
28	39	61.9	185	2 T30047	hypothetical prote
29	39	61.9	397	2 H70753	hypothetical prote

30	39	61.9	407	2 T51543	TOM (target of myb
31	39	61.9	509	1 ISMSSS	protein disulfide-
32	39	61.9	579	2 S69069	hypothetical prote
33	39	61.9	598	2 F83508	flagellar M-ring p
34	39	61.9	633	2 D83091	ATP sulfurylase GT
35	39	61.9	633	2 T14612	hypothetical prote
36	39	61.9	810	2 T19879	hypothetical prote
37	38	60.3	63	2 C87074	conserved hypothet
38	38	60.3	157	2 S09814	hypothetical prote
39	38	60.3	218	2 D83161	hypothetical prote
40	38	60.3	288	2 B40722	homeotic protein 1
41	38	60.3	360	2 B84376	hypothetical prote
42	38	60.3	476	1 C70986	probable serine/th
43	38	60.3	508	1 ISRTSS	protein disulfide-
44	38	60.3	523	2 A71246	lysine-tRNA ligase
45	38	60.3	539	2 T23700	hypothetical prote

ALIGNMENTS

RESULT 1

B70512
hypothetical protein Rv211c - Mycobacterium tuberculosis (strain H37Rv)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: B70512
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seegar, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: B70512
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-64 <COL>
A;Cross-references: UNIPROT:O33246; GB:297559; GB:AL123456; NID:g3261820; PIDN:CAB10702.
C;Genetics:
C;Experimental source: strain H37Rv
A;Gene: Rv211c

Query Match 76.2%; Score 48; DB 2; Length 64;
Best Local Similarity 80.0%; Pred. No. 0.42;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGSGDDDDLA 11
||| |||||:
Db 10 GGSGDDDDIA 19

RESULT 2

H84708
hypothetical protein At2g30480 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: H84708
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.W.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84708
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-240 <STO>
A;Cross-references: UNIPROT:O23686; GB:AE002093; NID:g2257712; PIDN:AAB63096.1; GSPDB:GN C;Genetics:
A;Gene: At2g30480
A;Map position: 2

```
Query Match 69.8%; Score 44; DB 2; Length 240;
Best Local Similarity 77.8%; Pred. No. 7.4;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGGSGDDDD 9
Db 163 VGGTGDDDD 171

RESULT 3
T27007
Hypothetical protein Y48B6A.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T27007
R:Wall, M.
Submitted to the EMBL Data Library, September 1999
A:Reference number: Z20297
A:Accession: T27007
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-894 <WIL>
A:Cross-references: UNIPROT:Q9U297; EMBL:AL110490; NID:el542263; PIDN:CAB54451.1; CESP:Y48B6A
A:Experimental source: clone Y48B6A
C:Genetics:
A:Gene: CESP:Y48B6A.11
A:Introns: 71/2; 146/2; 510/1; 558/2

Query Match 68.3%; Score 43; DB 2; Length 894;
Best Local Similarity 77.8%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGGDDDDL 10
Db 609 GGGGDDDDV 617

RESULT 4
S46825
Hypothetical protein YHL010c - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C:Accession: S46825
R:Favella, T.
Submitted to the EMBL Data Library, June 1994
A:Description: The sequence of S. cerevisiae cosmid 9433.
A:Reference number: S46796
A:Accession: S46825
A:Molecule type: DNA
A:Residues: 1-585 <FAV>
A:Cross-references: UNIPROT:P38748; EMBL:U11582; NID:g2289793; PID:g508747; GSPDB:GN000000
C:Genetics:
A:Gene: MIPS:YHL010c
A:Cross-references: SGD:S0001002
A:Map position: 8L
C:Superfamily: fission yeast probable zinc finger protein SPAC16E8.13; RING finger homology
F:236-285/Domain: RING finger homology <RNG>

Query Match 66.7%; Score 42; DB 2; Length 585;
Best Local Similarity 77.8%; Pred. No. 40;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGGSGDDDD 9
Db 373 VGGSGDDDN 381

RESULT 5
A53064
folded gastrulation (fog) precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: A53064
```

```
R:Costa, M.; Wilson, E.T.; Wieschaus, E.
Cell 76, 1075-1089, 1994
A:Title: A putative cell signal encoded by the folded gastrulation gene coordinates cell
A:Reference number: A53064; MUID:94185167; PMID:8137424
A:Accession: A53064
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-730 <COS>
A:Cross-references: UNIPROT:P40795; GB:U03717; NID:g430722; PIDN:AAA18955.1; PID:g430723
C:Genetics:
A:Gene: fog
A:Cross-references: FlyBase:FBgn0000719

Query Match 66.7%; Score 42; DB 2; Length 730;
Best Local Similarity 77.8%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGGSGDDDD 9
Db 513 IGGRGDDDD 521

RESULT 6
T02823
Probable membrane protein L1231.1 [imported] - Leishmania major (strain Friedlin)
C:Species: Leishmania major
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C:Accession: C81459; T02823
R:Wylter, P.J.; Audleman, L.; devos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.; F
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
A:Reference number: A81455; MUID:99178987; PMID:10077609
A:Accession: C81459
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-818 <PYL>
A:Cross-references: UNIPROT:Q25353; GB:AE001274; NID:g3264850; PIDN:AAC24646.1; PID:g3264
A:Experimental source: strain MHOM/IL/81/Friedlin
C:Genetics:
A:Gene: L1231.1
A:Map position: 1
C:Superfamily: Leishmania major probable membrane protein L1231.1
C:Keywords: transmembrane protein

Query Match 66.7%; Score 42; DB 2; Length 818;
Best Local Similarity 87.5%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGGDDDD 9
Db 711 GGGGDDDD 718

RESULT 7
S44769
C29E4.2 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 23-Mar-2001
C:Accession: S44769
R:Wilson, R.
Submitted to the EMBL Data Library, September 1993
A:Description: Sequence of the C. elegans cosmid C29E4.
A:Reference number: S44738
A:Accession: S44769
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-918 <WIL>
A:Cross-references: EMBL:L23651; NID:g388575; PID:g388580
C:Genetics:
A:Introns: 73/2; 97/3; 131/3; 381/3; 557/1; 684/1; 769/1; 830/3
C:Superfamily: Caenorhabditis elegans C29E4.2 protein

Query Match 66.7%; Score 42; DB 2; Length 918;
```


Best Local Similarity 87.5%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGGDDDD 9
||| |||||
Db 512 GGGGDDDD 519

RESULT 8
S54493
hypothetical protein YMR124w - Yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YMR564.06
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: S54493; S48514
R;Lye, G.; Church, C.M.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54014
A;Accession: S54493
A;Molecule type: DNA
A;Residues: 1-943 <LYE>
A;Cross-references: UNIPROT:P39523; EMBL:Z49273; NID:9809577; PIDN:CAA89273.1; PID:9809577
R;Pandit, S.; Sternglanz, R.
submitted to the EMBL Data Library, December 1992
A;Reference number: S48514
A;Accession: S48514
A;Molecule type: DNA
A;Residues: 770-943 <PAN>
A;Cross-references: EMBL:L07650; NID:g349191; PIDN:AAA35122.1; PID:g349192
C;Genetics:
A;Cross-references: SGD:S0004731
A;Map position: 13R

Query Match 66.7%; Score 42; DB 2; Length 943;
Best Local Similarity 87.5%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGGDDDD 9
||| |||||
Db 168 GGGGDDDD 175

RESULT 9
T18403
asparagine/aspartate rich protein - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18403
R;BARALE, J.C.; ATTALBONNEFOY, G.; BRAHIMI, K.; DASILVA, L.P.; LANGSLEY, G.
Mol. Biochem. Parasitol. 87, 169-181, 1997
A;Title: Plasmodium-falciparum asparagine and aspartate rich protein-2 is an evolutionary
A;Reference number: Z08834; MUID:97391121; PMID:9247928
A;Accession: T18403
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1360 <BAR>
A;Cross-references: UNIPROT:Q94649; EMBL:Y08924; NID:e1008106; PID:e1154303; PIDN:CAA7011
C;Genetics:
A;Gene: aarp2

Query Match 66.7%; Score 42; DB 2; Length 1360;
Best Local Similarity 87.5%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGGDDDD 9
||| |||||
Db 399 GGGGDDDD 406

RESULT 10
B87376
hypothetical protein CC1022 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus

Best Local Similarity 87.5%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGGDDDD 9
||| |||||
Db 512 GGGGDDDD 519

RESULT 8
S54493
hypothetical protein YMR124w - Yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YMR564.06
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: S54493; S48514
R;Lye, G.; Church, C.M.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54014
A;Accession: S54493
A;Molecule type: DNA
A;Residues: 1-943 <LYE>
A;Cross-references: UNIPROT:P39523; EMBL:Z49273; NID:9809577; PIDN:CAA89273.1; PID:9809577
R;Pandit, S.; Sternglanz, R.
submitted to the EMBL Data Library, December 1992
A;Reference number: S48514
A;Accession: S48514
A;Molecule type: DNA
A;Residues: 770-943 <PAN>
A;Cross-references: EMBL:L07650; NID:g349191; PIDN:AAA35122.1; PID:g349192
C;Genetics:
A;Cross-references: SGD:S0004731
A;Map position: 13R

Query Match 66.7%; Score 42; DB 2; Length 943;
Best Local Similarity 87.5%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGGDDDD 9
||| |||||
Db 168 GGGGDDDD 175

RESULT 9
T18403
asparagine/aspartate rich protein - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18403
R;BARALE, J.C.; ATTALBONNEFOY, G.; BRAHIMI, K.; DASILVA, L.P.; LANGSLEY, G.
Mol. Biochem. Parasitol. 87, 169-181, 1997
A;Title: Plasmodium-falciparum asparagine and aspartate rich protein-2 is an evolutionary
A;Reference number: Z08834; MUID:97391121; PMID:9247928
A;Accession: T18403
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1360 <BAR>
A;Cross-references: UNIPROT:Q94649; EMBL:Y08924; NID:e1008106; PID:e1154303; PIDN:CAA7011
C;Genetics:
A;Gene: aarp2

Query Match 66.7%; Score 42; DB 2; Length 1360;
Best Local Similarity 87.5%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGGDDDD 9
||| |||||
Db 399 GGGGDDDD 406

RESULT 10
B87376
hypothetical protein CC1022 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: B87376
R;Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: B87376
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-84 <STO>
A;Cross-references: UNIPROT:Q9A9G6; GB:AE005673; NID:gl3422314; PIDN:AAK23006.1; GSPDB:G
C;Genetics:
A;Gene: CC1022

Query Match 65.1%; Score 41; DB 2; Length 84;
Best Local Similarity 70.0%; Pred. No. 7.5;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGGDDDDLA 11
||| |||||
Db 35 GGGGDDDELA 44

RESULT 11
T36100
probable ATP-binding protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36100
R;Saunders, D.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1999
A;Reference number: Z21596
A;Accession: T36100
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-489 <SAU>
A;Cross-references: UNIPROT:Q9X873; EMBL:AL049661; PIDN:CAB41216.1; GSPDB:GN00070; SCOD
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCODB:SCE134.18

Query Match 65.1%; Score 41; DB 2; Length 489;
Best Local Similarity 66.7%; Pred. No. 48;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MGSGGDDDDLA 12
:|||||
Db 323 LGGGGDDTALVL 334

RESULT 12
ISHUSS
protein disulfide-isomerase (EC 5.3.4.1) precursor - human
N;Alternate names: cellular thyroid hormone-binding protein; endoplasmic reticulum prote
hydroxylase; proline,2-oxoglutarate 4-dioxygenase; prollyl 4-hydroxylase beta chain; pro
N;Contains: procollagen-proline dioxygenase (EC 1.14.11.2) beta chain; thyroxine deiodin
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1992 #sequence_revision 17-Nov-1995 #text_change 09-Jul-2004
C;Accession: A31913; B31913; A30055; A29787; A26632; A42791; I33178; F61002
R;Tasanen, K.; Parkkonen, T.; Chow, L.T.; Kivirikko, K.I.; Pihlajaniemi, T.
J. Biol. Chem. 263, 16218-16224, 1988
A;Title: Characterization of the human gene for a polypeptide that acts both as the beta
A;Reference number: A31913; MUID:89034087; PMID:2846539
A;Accession: A31913
A;Molecule type: DNA
A;Residues: 1-118;209-508 <TAS>
A;Cross-references: UNIPROT:P07237
A;Accession: B31913
A;Molecule type: DNA
A;Residues: 1-508 <TA2>
A;Cross-references: GB:J04049; GB:J04050; GB:M22803; GB:M22804; GB:M22805; GB:M22806; NI

R;Morris, J.I.; Varandani, P.T.
 Biochim. Biophys. Acta 949, 169-180, 1988
 A;Title: Characterization of a cDNA for human glutathione-insulin transhydrogenase (prot
 A;Reference number: A30055; MUID:88134925; PMID:3342239
 A;Accession: A30055
 A;Molecule type: mRNA
 A;Residues: 293-351; 'R', 353-401, 'HS', 404-459, 'Q', 461-508 <MOR>
 A;Cross-references: EMBL:X07077
 A;Experimental source: liver
 R;Cheng, S.; Gong, Q.; Parkison, C.; Robinson, E.A.; Appella, E.; Merlino, G.T.; Pastan,
 J. Biol. Chem. 262, 11221-11227, 1987
 A;Title: The nucleotide sequence of a human cellular thyroid hormone binding protein pre
 A;Reference number: A29787; MUID:87280213; PMID:3611107
 A;Accession: A29787
 A;Molecule type: mRNA
 A;Residues: 1-140, 'R', 142-359, 'RAG', 363-371, 'P', 373-508 <CH>
 A;Cross-references: GB:J02783; NID:G339646; PIDN:AAA61169.1; PID:G339647
 A;Note: parts of this sequence, including the amino end of the mature protein, were dete
 R;Pihlajaniemi, T.; Helaakoski, T.; Tasanen, K.; Myllyla, R.; Huhtala, M.L.; Koivu, J.;
 EMBO J. 6, 643-649, 1987
 A;Title: Molecular cloning of the beta-subunit of human prollyl 4-hydroxylase. This subun
 A;Reference number: A26632; MUID:87218523; PMID:3034602
 A;Accession: A26632
 A;Molecule type: mRNA
 A;Residues: 1-9, 'PWX', 13-43, 'PP', 46-48, 'H', 50-438, 'G', 440-443, 'G', 445-480, 'V', 482-508 <P
 A;Cross-references: EMBL:X05130; NID:G35654; PIDN:CAA28775.1; PID:G35655
 A;Note: parts of this sequence, including the amino end of the mature protein, were dete
 R;Tasanen, K.; Oikarinen, J.; Kivirikko, K.I.; Pihlajaniemi, T.
 J. Biol. Chem. 267, 11513-11519, 1992
 A;Title: Promoter of the gene for the multifunctional protein disulfide isomerase polype
 A;Reference number: A42791; MUID:92283868; PMID:1597478
 A;Accession: A42791
 A;Molecule type: DNA
 A;Residues: 1-25 <TA3>
 A;Cross-references: GB:S37207; NID:G249922; PIDN:AAB22262.1; PID:G249923
 A;Note: sequence extracted from NCBI backbone (NCBIN:104755, NCBIIP:104756)
 R;Ward, L.D.; Hong, J.; Whitehead, R.H.; Simpson, R.J.
 Electrophoresis 11, 883-891, 1990
 A;Title: Development of a database of amino acid sequences for human colon carcinoma pro
 A;Reference number: A33178; MUID:91176935; PMID:2079031
 A;Accession: I33178
 A;Molecule type: protein
 A;Residues: 18-24, 'X', 26 <WAR>
 R;Bauw, G.; Rasmussen, H.H.; Van Den Bulcke, M.; Van Damme, J.; Puype, M.; Gesser, B.; C
 Electrophoresis 11, 528-536, 1990
 A;Title: Two-dimensional gel electrophoresis, protein electroblotting and microsequencin
 A;Reference number: A61002; MUID:91031404; PMID:1699755
 A;Accession: F61002
 A;Molecule type: protein
 A;Residues: 317-326; 351-363, 'X', 365-370; 402-406, 'X', 408-419 <BAU>
 C;Genetics:
 A;Gene: GDB:P4HB; PO4DB
 A;Cross-references: GDB:120708; OMIM:116790
 A;Map position: 17q25-17q25
 A;Introns: 49/1; 118/1; 162/3; 208/3; 243/3; 285/3; 352/3; 393/1; 453/3; 482/3
 C;Complex: in procollagen-proline dioxygenase forms a heterotetramer of two alpha chains
 C;Function: <DSI>
 A;Description: as protein disulfide-isomerase (EC 5.3.4.1) catalyzes the rearrangement o
 A;Pathway: protein synthesis
 C;Function: <PHB>
 A;Description: as procollagen-proline dioxygenase (EC 1.14.11.2) beta chain participates
 A;Pathway: collagen synthesis
 C;Function: <TDI>
 A;Description: as thyroxine deiodinase (EC 3.8.1.4) reduces iodinated thyronine residues
 C;Superfamily: protein disulfide-isomerase; thioredoxin homology
 C;Keywords: ascorbic acid; duplication; endoplasmic reticulum; homodimer; intramolecular
 F;1-17/Domain: signal sequence #status predicted <SIG>
 F;18-508/Product: protein disulfide-isomerase #status predicted <MAT>
 F;32-119/Domain: thioredoxin homology <TX1>
 F;375-460/Domain: thioredoxin homology <TX2>
 F;505-508/Region: endoplasmic reticulum retention signal
 F;53-56, 397-400/Disulfide bonds: redox-active #status predicted

F;312-343/Disulfide bonds: #status predicted
 Query Match 65.1%; Score 41; DB 1; Length 508;
 Best Local Similarity 87.5%; Pred. No. 50;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GSGDDDDL 10
 Db 477 GAGDDDDL 484
 RESULT 13
 ISBOSS
 protein disulfide-isomerase (EC 5.3.4.1) precursor - bovine
 N;Alternate names: cellular thyroid hormone-binding protein; endoplasmic reticulum protei
 N;Contains: procollagen-proline dioxygenase (EC 1.14.11.2) beta chain; thyroxine deiodina
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
 C;Accession: A26829
 R;Yamauchi, K.; Yamamoto, T.; Hayashi, H.; Koya, S.; Takikawa, H.; Toyoshima, K.; Horiuc
 Biochem. Biophys. Res. Commun. 146, 1485-1492, 1987
 A;Title: Sequence of membrane-associated thyroid hormone binding protein from bovine live
 A;Reference number: A26829; MUID:87298601; PMID:3619939
 A;Accession: A26829
 A;Molecule type: mRNA
 A;Residues: 1-510 <YAM>
 A;Cross-references: UNIPROT:P05307; GB:M17596; NID:G163496; PIDN:AAA30690.1; PID:G163497
 A;Note: the authors translated the codon TTC for residue 175 as Leu
 C;Comment: This enzyme catalyzes the rearrangement of both intrachain and interchain dis
 C;Superfamily: protein disulfide-isomerase; thioredoxin homology
 C;Keywords: ascorbic acid; duplication; endoplasmic reticulum; homodimer; intramolecular
 F;1-19/Domain: signal sequence #status predicted <SIG>
 F;20-510/Product: protein disulfide-isomerase #status predicted <MAT>
 F;34-121/Domain: thioredoxin homology <TX1>
 F;377-462/Domain: thioredoxin homology <TX2>
 F;507-510/Region: endoplasmic reticulum retention signal
 F;55-58, 399-402/Disulfide bonds: redox-active #status predicted
 F;314-345/Disulfide bonds: #status predicted
 Query Match 65.1%; Score 41; DB 1; Length 510;
 Best Local Similarity 87.5%; Pred. No. 50;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GSGDDDDL 10
 Db 479 GAGDDDDL 486
 RESULT 14
 T46169
 CDC48-like protein - Arabidopsis thaliana
 N;Alternate names: protein T4D2.160
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
 C;Accession: T46169
 R;Nyakatura, G.; Fattmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.; F
 submitted to the Protein Sequence Database, December 1999
 A;Reference number: Z23025
 A;Accession: T46169
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-815 <NYA>
 A;Cross-references: UNIPROT:Q9SCN8; EMBL:AL132958
 A;Experimental source: cultivar Columbia; BAC clone T4D2
 C;Genetics:
 A;Map position: 3
 A;Introns: 10/2; 65/3; 142/1; 397/3; 589/3
 A;Note: T4D2.160
 C;Superfamily: transitional endoplasmic reticulum ATPase; FtsH/SEC18/CDC48-type ATP-bind:
 Query Match 65.1%; Score 41; DB 2; Length 815;
 Best Local Similarity 77.8%; Pred. No. 82;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSGDDDDL 10
 ||: |||||
 Db 805 GGAADDDL 813

RESULT 15

S66672
 Phosphatidylinositol-specific phospholipase C - northern European squid
 C;Species: Loligo forbesi (northern European squid)
 C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004
 C;Accession: S66672; S74310; S41169
 R;Carne, A.; McGregor, R.A.; Bhatia, J.; Sivaprasadarao, A.; Keen, J.N.; Davies, A.; Fin
 PEBS Lett. 372, 243-248, 1995
 A;Title: A beta-subclass phosphatidylinositol-specific phospholipase C from squid (Loligo
 A;Reference number: S66672; MUID:96000215; PMID:7556677
 A;Accession: S66672
 A;Molecule type: mRNA
 A;Residues: 1-875 <CAR>
 A;Cross-references: UNIPROT:Q25285; EMBL:X76968; NID:9439518; PIDN:CAA54275.1; PID:94395
 A;Note: the authors translated the codon AAA for residue 86 as Asp and AAT for residue 1
 A;Accession: S74310
 A;Molecule type: protein
 A;Residues: 90-102:278-297:756-770 <CAN>
 C;Superfamily: Phosphatidylinositol-specific Phospholipase C; 1-phosphatidylinositol-4,5
 main Y homology
 F;325-471/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain X hom
 F;513-633/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain Y hom

Query Match 65.1%; Score 41; DB 1; Length 875;
 Best Local Similarity 77.8%; Pred.No. 88;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 GSGDDDDL 11
 ||: |||||
 Db 476 GAGDDELA 484

Search completed: May 27, 2005, 13:48:41
 Job time : 3.92111 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:35:35 ; Search time 0.921109 Seconds
(without alignments)
972.511 Million cell updates/sec

Title: US-10-622-893a-1
Perfect score: 63
Sequence: 1 MGSGDDDDDL 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	69.8	265	4	US-09-248-796A-20812
2	43	68.3	408	4	US-09-248-796A-19565
3	42	66.7	151	4	US-09-270-767-39203
4	42	66.7	151	4	US-09-270-767-54420
5	42	66.7	264	4	US-09-248-796A-27609
6	42	66.7	409	4	US-09-540-236-3384
7	42	66.7	543	4	US-09-248-796A-22504
8	42	66.7	1138	4	US-09-252-991A-25952
9	41	65.1	390	4	US-09-949-016-11598
10	41	65.1	445	4	US-09-252-991A-30715
11	41	65.1	491	1	US-07-872-673B-3
12	41	65.1	508	4	US-09-807-258-30
13	41	65.1	509	2	US-08-557-122A-29
14	41	65.1	509	3	US-09-262-666-29
15	41	65.1	510	2	US-08-557-122A-28
16	41	65.1	510	2	US-08-557-122A-36
17	41	65.1	510	3	US-09-262-666-28
18	41	65.1	510	3	US-09-262-666-36
19	41	65.1	675	4	US-09-252-991A-27987
20	41	65.1	3052	2	US-08-557-122A-26
21	41	65.1	3052	3	US-09-262-666-26
22	40	63.5	10	4	US-08-969-334-1
23	40	63.5	11	1	US-08-200-900A-37
24	40	63.5	11	4	US-08-794-042-37
25	40	63.5	23	3	US-09-217-228-3
26	40	63.5	106	4	US-09-902-540-15731
27	40	63.5	114	3	US-09-217-228-5

```

28 40 63.5 134 3 US-09-272-342B-5 Sequence 5, Appli
29 40 63.5 154 3 US-09-439-313-383 Sequence 383, App
30 40 63.5 154 3 US-09-352-616A-383 Sequence 383, App
31 40 63.5 154 4 US-09-636-215-383 Sequence 383, App
32 40 63.5 154 4 US-09-685-166A-383 Sequence 383, App
33 40 63.5 154 4 US-09-679-426-383 Sequence 383, App
34 40 63.5 154 4 US-09-759-143-383 Sequence 383, App
35 40 63.5 154 4 US-09-651-236-383 Sequence 383, App
36 40 63.5 161 4 US-09-636-215-846 Sequence 846, App
37 40 63.5 161 4 US-09-685-166A-846 Sequence 846, App
38 40 63.5 161 4 US-09-679-426-846 Sequence 846, App
39 40 63.5 161 4 US-09-759-143-846 Sequence 846, App
40 40 63.5 161 4 US-09-651-236-846 Sequence 846, App
41 40 63.5 177 3 US-09-272-342B-6 Sequence 6, Appli
42 40 63.5 290 4 US-09-248-796A-14651 Sequence 14651, A
43 40 63.5 291 3 US-09-328-869-2 Sequence 2, Appli
44 40 63.5 291 3 US-09-629-774A-2 Sequence 2, Appli
45 40 63.5 296 1 US-07-745-382-14 Sequence 14, Appli

```

ALIGNMENTS

RESULT 1
US-09-248-796A-20812
; Sequence 20812, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20812
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20812

Query Match 69.8% Score 44; DB 4; Length 265;
Best Local Similarity 77.8%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Qy      2 GSGGDDDDL 10
Db      202 GGGGDDDDI 210

```

RESULT 2
US-09-248-796A-19565
; Sequence 19565, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19565
; LENGTH: 408
; TYPE: PRT

```
; ORGANISM: Candida albicans
US-09-248-796A-19565

Query Match      68.1%; Score 43; DB 4; Length 408;
Best Local Similarity 72.7%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 GGGSGDDDDLAL 12
      || ||||| |
Db      386 GGGGDDDDNGL 396

RESULT 3
US-09-270-767-39203
; Sequence 39203, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39203
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39203

Query Match      66.7%; Score 42; DB 4; Length 151;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GGGSGDDDD 9
      || |||||
Db      68 GGGGDDDD 75

RESULT 4
US-09-270-767-54420
; Sequence 54420, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54420
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54420

Query Match      66.7%; Score 42; DB 4; Length 151;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GGGSGDDDD 9
      || |||||
Db      68 GGGGDDDD 75

RESULT 5
US-09-248-796A-27609
; Sequence 27609, Application US/09248796A
```

```
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 27609
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-27609

Query Match      66.7%; Score 42; DB 4; Length 264;
Best Local Similarity 87.5%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GGGSGDDDD 9
      || |||||
Db      222 GGGGDDDD 229

RESULT 6
US-09-540-236-3384
; Sequence 3384, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3384
; LENGTH: 409
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-3384

Query Match      66.7%; Score 42; DB 4; Length 409;
Best Local Similarity 87.5%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GGGSGDDDD 9
      || |||||
Db      386 GGGGDDDD 393

RESULT 7
US-09-248-796A-22504
; Sequence 22504, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22504
; LENGTH: 543
```

; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-22504

Query Match 66.7%; Score 42; DB 4; Length 543;
Best Local Similarity 87.5%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GSGGDDDD 9
Db 298 GGGGDDDD 305

RESULT 8
US-09-252-991A-25952
; Sequence 25952, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25952
; LENGTH: 1138
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25952

Query Match 66.7%; Score 42; DB 4; Length 1138;
Best Local Similarity 63.6%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GSGGDDDDLAI 12
Db 303 GGQDEDLAI 313

RESULT 9
US-09-949-016-11598
; Sequence 11598, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 11598
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11598

Query Match 65.1%; Score 41; DB 4; Length 390;
Best Local Similarity 87.5%; Pred. No. 69;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GSGDDDDL 10
Db 359 GAGDDDDL 366

RESULT 10
US-09-252-991A-30715
; Sequence 30715, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30715
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30715

Query Match 65.1%; Score 41; DB 4; Length 445;
Best Local Similarity 58.3%; Pred. No. 79;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MGGSGDDDDLAI 12
Db 190 LAGSGADDDIAV 201

RESULT 11
US-07-872-673B-3
; Sequence 3, Application US/07872673B
; Patent No. 5578466
; GENERAL INFORMATION:
; APPLICANT: Toshiya HAYANO, Setsuko KATO, No. 5578466uhiro TAKAHASHI, and Masanori
; TITLE OF INVENTION: Co-expression System of Protein Disulfide Isomerase Gene an
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Type 2DD, 3.50 inch, 720 KB
; COMPUTER: Apple Macintosh SE
; OPERATING SYSTEM: Apple DOS
; SOFTWARE: Microsoft Word Version 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,673B
; FILING DATE: 19920417
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Japanese Patent Application No. 5578466. 114074/91 and 31160
; FILING DATE: 18-APR-1991 and 30-OCT-1991
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: protein
US-07-872-673B-3

Query Match 65.1%; Score 41; DB 1; Length 491;

[illegible]

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: No. 58796640 No. 5879664disk of No. 5879664th America, Inc.
;; STREET: 405 Lexington Avenue, 64th Floor
;; CITY: New York
;; STATE: New York
;; COUNTRY: United States of America
;; ZIP: 10174-6401
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/557,122A
;; FILING DATE: 11-DEC-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lambiris, Elias J.
;; REGISTRATION NUMBER: 33,728
;; REFERENCE/DOCKET NUMBER: 3980.204-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-867-0123
;; TELEFAX: 212-878-9655
;; INFORMATION FOR SEQ ID NO: 28:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 510 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-557-122A-28

Query Match 65.1%; Score 41; DB 2; Length 510;
Best Local Similarity 87.5%; Pred. No. 91;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 GSGDDDDL 10
|:|||||
Db 479 GAGDDDDL 486

Search completed: May 27, 2005, 13:49:57
Job time : 1.92111 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:28:55 ; Search time 95.9723 Seconds
(without alignments)
2331.700 Million cell updates/sec

Title: US-10-622-893A-3
Perfect score: 2329
Sequence: 1 AVTKSSLLIVGAGTGTST.....MDFHDVKEWTVQYRDISKL 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2329	100.0	438	2 P78573	P78573 aspergillus
2	1981.5	85.1	437	2 Q92204	Q92204 aspergillus
3	1242	53.3	445	2 O42629	O42629 aspergillus
4	848	36.4	426	2 O6B2B0	O6B2B0 debaryomyce
5	846.5	36.3	418	2 O6B2A7	O6B2A7 debaryomyce
6	660	28.3	437	2 Q92261	Q92261 penicillium
7	629	27.0	438	2 Q96074	Q96074 emericella
8	625	26.8	437	2 Q765A9	Q765A9 eupenicilli
9	600	25.8	437	2 Q765A8	Q765A8 coniochaeta
10	488.5	21.0	412	2 O43029	O43029 schizosacch
11	465	20.0	433	2 Q9UTW9	Q9UTW9 schizosacch
12	426.5	18.3	502	2 Q78638	Q78638 neurospora
13	386	16.6	433	2 Q6CL10	Q6CL10 kluyveromyc
14	376.5	16.2	486	2 Q6BN43	Q6BN43 debaryomyce
15	351	15.1	446	2 Q6BN42	Q6BN42 debaryomyce
16	346	14.9	435	2 Q9P976	Q9P976 cylindrocarr
17	192	8.2	390	1 MSOX_BACB0	P40859 bacillus sp
18	185	7.9	387	2 Q6ITC6	Q6ITC6 bacillus sp
19	183	7.9	371	2 Q737A8	Q737A8 bacillus ce
20	181.5	7.8	440	2 Q9V985	Q9V985 drosophila
21	181	7.8	384	1 MSOX_STRAW	Q827H4 streptomyce
22	181	7.8	388	1 MSOX_STRSB	P40854 streptomyce
23	177.5	7.6	390	1 SOX_HUMAN	Q9P029 homo sapien
24	177.5	7.6	390	2 Q6IAJ9	Q6IAJ9 homo sapien
25	175	7.5	371	2 Q63AM4	Q63AM4 bacillus ce
26	174	7.5	371	2 Q81PT2	Q81PT2 bacillus an
27	172.5	7.4	395	2 Q92NL2	Q92NL2 rhizobium m
28	172	7.4	367	2 Q982V8	Q982V8 rhizobium l
29	170.5	7.3	773	2 Q66IN9	Q66IN9 xenopus lae
30	170	7.3	371	2 Q81CNO	Q81CNO bacillus th
31	170	7.3	371	2 Q6HI11	Q6HI11 bacillus th

ALIGNMENTS

RESULT 1

```
P78573 ID P78573 PRELIMINARY; PRT; 438 AA.
AC P78573;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fructosyl amine: oxygen oxidoreductase.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5085;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97284723; PubMed=9139700; DOI=10.1074/jbc.272.19.12505;
RA Takahashi M., Pischetrieder M., Monnier V.M.;
RT "Molecular cloning and expression of ananase isoenzyme (fructosyl
aminoxidoreductase, EC 1.5.3) from Aspergillus fumigatus.";
RL J. Biol. Chem. 272:12505-12507(1997).
DR EMBL; U82830; AAC49711.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR006076; FAD oxidred.
DR InterPro; IPR000205; NAD_BS.
DR InterPro; IPR000594; ThiF_domain.
DR Pfam; PF01266; DAO; 1.
SQ SEQUENCE 438 AA; 48931 MW; FBC8A3ESD89D02AE CRC64;
```

Query Match 100.0%; Score 2329; DB 2; Length 438;
Best Local Similarity 100.0%; Pred. No. 1.9e-174; Indels 0; Gaps 0;
Matches 437; Conservative 0; Mismatches 0

```
Qy 1 AVTKSSLLIVGAGTGTCTALHLARRGYTNVTVDLPYPVPSAISAGNDVNVKVISGQYS 60
Db 2 AVTKSSLLIVGAGTGTCTALHLARRGYTNVTVDLPYPVPSAISAGNDVNVKVISGQYS 61
Qy 61 NNKDEIEVNEILAEAFNGWKNDPLFKPYTHDTGLLSACSQEGDLRLGVVRGEPDNL 120
Db 62 NNKDEIEVNEILAEAFNGWKNDPLFKPYTHDTGLLSACSQEGDLRLGVVRGEPDNL 121
Qy 121 VELTRPQFKLAPEGVLQGFPGWKGYFARSGAGWAHARNALVAAAREAQRMGVKFTVG 180
Db 122 VELTRPQFKLAPEGVLQGFPGWKGYFARSGAGWAHARNALVAAAREAQRMGVKFTVG 181
Qy 181 TPQGRVVTLIFENNNDKVGAVTGDKIKIWRARTFLCAGASAGQFLDFKNQLRPTATLVIH 240
Db 182 TPQGRVVTLIFENNNDKVGAVTGDKIKIWRARTFLCAGASAGQFLDFKNQLRPTATLVIH 241
Qy 241 ALKPEERALKYNIPIVNIERGFPPFEPDEERGEIKICDEHPGYTNMVQSADGTWMSIPPE 300
Db 242 ALKPEERALKYNIPIVNIERGFPPFEPDEERGEIKICDEHPGYTNMVQSADGTWMSIPPE 301
Qy 301 KTOIPKEAEFVRVALLKETMPQLADRPFSFARICWCADTANREFLIDRHPQYHSLVLCGC 360
Db 302 KTOIPKEAEFVRVALLKETMPQLADRPFSFARICWCADTANREFLIDRHPQYHSLVLCGC 361
```

Q95U69 drosophila
P79371 oryctolagus
P40873 arthrobacte
Q8XTY8 ralatonia s
Q88GE9 pseudomonas
Q8N6Z6 homo sapien
Q8D826 mus musculu
Q46904 escherichia
Q82M71 streptomyce
Q89M92 bradyrhizob
Q6R512 bacillus sp
Q8EMP0 oceanobacil
P23342 bacillus sp
Q8FEI5 escherichia

```
QY 361 ASGRGFKYLPISIGNLIVDAMEGKVPQKIHELIIKKNPDIANRNWRDTLGRFGGPNRVMD 420
Db 362 ASGRGFKYLPISIGNLIVDAMEGKVPQKIHELIIKKNPDIANRNWRDTLGRFGGPNRVMD 421
QY 421 HDVKWNTNQYRDISKL 437
Db 422 HDVKWNTNQYRDISKL 438

RESULT 2
Q92204 PRELIMINARY; PRT; 437 AA.
AC Q92204;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fructosyl amino acid oxidase.
GN Names=fa0A;
OS Aspergillus terreus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=331178;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97175014; PubMed=9022674; Fukuva H., Yagi M., Tani Y., Kato N.,
RA Yoshida N., Sakai Y., Isogai A., Fukuva H., Yagi M., Tani Y., Kato N.;
RT "Primary structures of fungal fructosyl amino acid oxidases important
RL Eur. J. Biochem. 242:499-505(1996).";
DR EMBL; Y09020; CAA70218.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR Pfam; PF01266; DAO; 1.
SQ SEQUENCE 437 AA; 48669 MW; B838EA0AB2105E36 CRC64;
```

```
Query Match 85.1%; Score 1981.5; DB 2; Length 437;
Best Local Similarity 82.6%; Pred. No. 3.8e-147;
Matches 360; Conservative 43; Mismatches 32; Indels 1; Gaps 1;

QY 2 VTKSSLLIVGAGTGTSTALHLARGYTNVTLDPPVPSPSAISAGNDVNVKVISSGOYSN 61
Db 3 VTKSSLLIVGAGTGTSTALHLARGYTNVTLDPPVPSPSAISAGNDVNVKVISSGOYS 62

QY 62 NKEIEVNEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQEGLDRLGV-RVRPG 121
Db 63 KKEVEVNEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQEGLDRLGV-RVRPG 122

QY 122 ELTRPQFRKLAPGVLQDFFQWKGYFARSGAGWAHARNALVAAAREAQRMGVKFTGT 181
Db 123 ELTRPQFRKLAP-GVLKGNFPGMRGVHTRSNAGWAHARNALVAAAREAQRLGVRFVAGS 181

QY 182 PQGRVVTILFENNNDVKGAVTGDKIWRABETFLCAGASAGQFLDFKNQLRPTAVLHIA 241
Db 182 PQGRVVTILFENNNDVKGAVTADGKIWRABETFLCAGASAGQFLDFKNQLRPTAVLHIO 241

QY 242 LKPEERALYKNIPIENIERGFFFEDEERGEIKICDEHPGYNTMVQSDGTWMSIPFEK 301
Db 242 LKPEERALYKNIPIENIERGFFFEDEERGEIKICDEHPGYNTMTGADGRVRSIPFEK 301

QY 302 TQIPKEATRVRLAKETMPQLADRPFSFARICWCADTANREFLIDRHHPQYHSLVLCGGA 361
Db 302 TQVPREAEVRKLLSETMPQLADRPFSFARICWCADTNPREFIIDRHHPYPSVLVLCGGA 361

QY 362 SGRGFKYLPISIGNLIVDAMEGKVPQKIHELIIKKNPDIANRNWRDTLGRFGGPNRVMD 421
Db 362 SGRGFKYLPISIGIIADAMEDKTPAKIHLIRWSPFIARNWGDRLGRFGGPNRVMDFN 421

QY 422 DVKEWNTNQYRDISKL 437
Db 422 EVKEWNTNQYRDISKL 437
```

RESULT 3

```
042629 PRELIMINARY; PRT; 445 AA.
AC 042629;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fructosyl amine:oxygen oxidoreductase.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5085;
RN [1]
RP SEQUENCE FROM N.A.
RA Takahashi M., Pischetsrieder M., Thein M., Roth J.C., Mommier V.M.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF035700; AAB88209.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR006076; Fad_oxred.
DR InterPro; IPR000205; NAD_BS.
DR InterPro; IPR000594; Thif_domain.
DR Pfam; PF01266; DAO; 1.
SQ SEQUENCE 445 AA; 49364 MW; 57685EB6CA535B23 CRC64;

Query Match 53.3%; Score 1242; DB 2; Length 445;
Best Local Similarity 52.1%; Pred. No. 5.1e-89;
Matches 228; Conservative 78; Mismatches 118; Indels 14; Gaps 6;

QY 2 VTKSSLLIVGAGTGTSTALHLARGYTNVTLDPPVPSPSAISAGNDVNVKVISSGOYSN 61
Db 6 LSTESSIIIVIGAGTGTSTALHLARGYKDVTLDPHPVPSPPIAGNDINKIM---EHSE 62

QY 62 NKD-----EIEVNEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQEGLDRLGV-RVRPG 115
Db 63 LKDGSDPRSAAFSTFTRAAKXAKTDPVFQPYFHTGTGFIISGHTPALIDHIREKDEVEPS 122

QY 116 EDNLVELTRPQFRKLAPGVLQDFFQWKGYFARSGAGWAHARNALVAAAREAQRMGV 175
Db 123 E-TNFVKLETAEDFRRTMPPGVLTGDPGKWKGLHKSAGWIHAKKAMISAFNEAKRLGV 181

QY 176 KFTVTGTPQGRVVTILFENNNDVKGAVTGDKIWRABETFLCAGASAGQFLDFKNQLRPTAW 235
Db 182 RVTGSGPEGNVSLVYEDGDVVGARTADGRVHKAHRTILSAGASGSLDFKKQLRPTAW 241

QY 236 TLVHTALKEERALYKNIPIENIERGFFFEDEERGEIKICDEHPGYNTMVQSDA--GT 293
Db 242 TLCHIQMGFEVVKYENLPVLFNIAKGFPMFEDEDKHELKICDEHPGYCNFLPDPNRPQ 301

QY 294 MMSIPFEKTQIPKEATRVRLAKETMPQLADRPFSFARICWCADTANREFLIDRHHPQYH 353
Db 302 EKSVPPFAKHQIPLAEARARDFLHDTMPLHADRPLSFARICWDADTPDRAFLIDRHPEHP 361

QY 354 SLVLGCGASGRGFKYLPISIGNLIVDAMEGKVPQKIHELIIKKNPDIANRNWRDTLGRFGG 413
Db 362 SLLVAVGGSGNGAMQMTPIGGFIADALESKLOKEVKDIIVRWRPETAVDRDWRATQNRFGG 421

QY 414 PNRVMDFHDKVE--WTNV 429
Db 422 PDRIMDFQVQGEDQWTKI 439

RESULT 4
Q6BZB0 PRELIMINARY; PRT; 426 AA.
AC Q6BZB0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to triP78573 Aspergillus fumigatus Fructosyl amine: oxygen
DE oxidoreductase.
GN ORFNames=DEHA0A03157g;
OS Debaryomyces hansenii CBS767.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
```

```
OX NCBI_TaxID=284592;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confaniolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekaiia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44 (2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genolevures;
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RA EMBL: CR382133; CAG84414.1; -.
DR InterPro; IPR006076; Fad_oxred.
DR InterPro; IPR000205; NAD_BS.
DR Pfam; PF01266; DAO; 1.
SQ SEQUENCE 426 AA; 46297 MW; 7A5A382DF1P72293 CRC64;

Query Match 36.4%; Score 848; DB 2; Length 426;
Best Local Similarity 41.2%; Pred. No. 4.4e-58;
Matches 181; Conservative 71; Mismatches 153; Indels 34; Gaps 10;

QY 4 KSSSLIVGAGTGTSTALHARRGYTNVTLDPYPVPSAISAGNDVNKVISSQ----Y 59
Db 3 KPGKILITAGTGLSTALHLLKQGEKDVILVDPYAVPSPFSGNDVNKIIQTSSDDFY 62
QY 60 SNNKDEIVNEILAEAFNGKNDPLFKPYHYDTGLLMSACSQEGLDRLGVRVR--PGED 117
Db 63 SK-----LAEALEMWRDNVFNKAFETGIIYATGKQESIDYRYEYLLGRK 112
QY 118 PNLVELTRPQFRKLAP--EGV--LQDFFGKMGYFARSAGAGWAHARNALVAAREAQR 173
Db 113 DKVVKLNSVEDYKYPVNEKGLKSPNFKQWYGYOEKNCGWAFARLALENCVECKL 172
QY 174 GVKFVTGTPQGRVTLIP--ENNVDKGVATGDKIWAERTFLCAGASAGOFDFKQLRP 232
Db 173 GAKFVIDSAB-----ELFSEDGACVGHVSNGNIIEADRTIICAGANSFKFLNFEQQLA 228
QY 233 TAWTLVHIALKPEERALKYKNIPIVFNTERGFFPEPDEERGEIKICDHPGYTNVQSADG 292
Db 229 KCTLGHILKLTDEALLKMPVNLNLDGCFVFEFD--LNNEIKFCNEFPGYVNVNED-- 285
QY 293 TMSIPIPEKTQIPKEATRVRALLKETMPQADRPFSFARICWCADTANREFLIDRHPQY 352
Db 286 ---SVPSFKDSIPKEAEDOMKAFRLQVFPFAEREPFLARICWCTDTPDRHFLICEHPGH 342
QY 353 HSLVLGCGAGRGFKYLPSTGNLIVDM---EGKVPQKIHILKWNPDIAANNRWDTLG 409
Db 343 KNLVLGTGDSGGQFKYMPNPGKYISQVALKGNSLDKDKELWRWRPDMGKGRDLKDLQ 402
QY 410 RFGGPNRVMDFHDVKEWTN 428
Db 403 RYGSNEVKDLKNVQKWSN 421

RESULT 5
Q6BZA7 PRELIMINARY; PRT; 418 AA.
ID Q6BZA7
AC Q6BZA7;
```



```

QY 281 PGYTNMQS---ADGTMWSIPP---EKTQTPKEATRVALLKETMPQLADRPSPFARIC 334
Db 285 YEYVCNVETDYDENSQVSVSPHSGPSKSLPKYAIQWRRFLDTFLPDLDRLSINTKMC 344
QY 335 WCADTANREFLDRHPQYHSLVLCGASGSGFKPLPSIGNLNVDMARGKVPQKHILIKW 394
Db 345 WISDTEANFLDKVPQFDNVFVANGDSGHAFKPLNIGRYIAQRIIGLDSSEKWDARW 404
QY 395 NPDIANR-NWR 405
Db 405 REDDKASELXWR 416

RESULT 12
QY 7S638 PRELIMINARY; PRT; 502 AA.
AC 07S638;
DT 01-MAR-2004 (TremBrel. 26, Created)
DT 25-OCT-2004 (TremBrel. 26, Last sequence update)
DE Hypothetical protein (Related to fructosyl amino acid oxidase).
GN Name=NCU04771.1; Synonyms=B15810.150;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins D., Engels S., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kanvysselis M., Mauceli E., Bielek C., Rudd S., Frisman D.,
RA Kryzofova S., Rasmussen C., Merzenberg R.B., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
RL Nature 0:0-0(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Partmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AABX01000331; EAA30984.1; -.
DR EMBL; BX897677; CA85581.1; -.
DR InterPro; IPR006076; FAD_oxred.
DR Pfam; PF01266; DAO; 1.
KW Hypothetical protein.
SQ SEQUENCE 502 AA; 52930 MW; 39748F9B553D4FDE CRC64;

```

```

Query Match 18.3%; Score 426.5; DB 25; Length 502;
Best Local Similarity 29.2%; Pred. No. 7.3e-25;
Matches 130; Conservative 56; Mismatches 169; Indels 109; Gaps 21;

QY 6 SLLIIVGAGTGWGSTALHLARR-GYTNVTVL-----DPVPVPSAISAGNDVNKVI 54
Db 30 SLLIIVGAGTGWGSTALHLARR-GYTNVTVL-----DPVPVPSAISAGNDVNKVI 54

```

```

QY 55 SSGOYNNKDEIEVNEILAEAFNGWK-----NDPLFKPYVYHDTGLLMSACSQSG-LDRL 108
Db 88 RA-DYPD-----AAYAALAAQAQLQWRQOHTHDPDLGSEGRYSGLLWA-----DGPAPPP 138
QY 109 GVRVRPG-----EDPNLV-----ELTRPEQFRKLAPGV 137
Db 139 GTPVAPGTSVIDKSLKTGM DYARPSWANVLSSASSDPELAARIKECPNTDAIKALGTG- 197
QY 138 LQGDFFPGKGYFARGAGWAHARNALVAAAREARMG-VKPVTTGTPGVRVTLTFENNDV 196
Db 198 --GSSGSW-GYI-NGLSGWANAGASMAWLYKRVRAEGRINFVA---GEVTNLEVSNGTV 249
QY 197 KGAVTGDKIWRARTFLCAGASAGQFLDFKNQLRPTAATLTVHIALKEEERALKYNTPI 256
Db 250 TGAFLSGRVLSDALVNVWSAGAWTGRVLDTGTQAIATGQVLGYIDLTFEEAQLAHMPVI 309
QY 257 FNIERGFPPFPDBERGEIKICDEHPGYTNMVQSGADGTMMSIP-----FEKT----- 302
Db 310 LNLSTGLFVIP-PRNGVLKVARHAYGYLN-----PTTSLVPPPLASFVTTTAAAVSLPLT 362
QY 303 -----QIPKEATRVRLAKE--TWPQLADRPSPFARICWCADTANRFLIDRHPOYH 353
Db 363 TLTDPQLIPTEGADDLRRALHEMVPLPSLRDRPFSKTRICWTYSDTPTADFIVDYHPKYK 422
QY 354 SLVLGCGASGRGKYLPSIGNLNVDMARGKVPQKHILIKWNPDIANRNR 405
Db 423 GLFVATGDSGHAFKPLFVIGEKIADVIAGQCP-----PEFVGKWNWR 464

RESULT 13
QY 06CLIO PRELIMINARY; PRT; 433 AA.
AC 06CLIO;
DT 25-OCT-2004 (TremBrel. 28, Created)
DT 25-OCT-2004 (TremBrel. 28, Last sequence update)
DT 25-OCT-2004 (TremBrel. 28, Last annotation update)
DE Similar to sp|Q9UTM9 Schizosaccharomyces pombe Putative fructosyl
DE amino acid oxidase.
GN ORFNames=KLIA0F02860g;
OS Kluyveromyces lactis NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetaceae; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neugeglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellens S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swenne D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts."
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382126; CAG97917.1; -.
DR InterPro; IPR006076; FAD_oxred.
DR InterPro; IPR000205; NAD_BS.
DR Pfam; PF01266; DAO; 1.
SQ SEQUENCE 433 AA; 48905 MW; 1091A3C6C3F3DA08 CRC64;

```

RP	SEQUENCE FROM N.A.				
RC	STRAIN=CBS767;				
RA	Genoscope;				
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; CR382138; CAG88679.1; --				
DR	GO; GO:0003824; F:catalytic activity; IEA.				
DR	InterPro; IPR006076; Fad_Oxred.				
DR	InterPro; IPR000205; Nad_BS.				
DR	InterPro; IPR000408; Reg_chromatid condens.				
DR	InterPro; IPR000594; Thif domain.				
DR	Pfam; PF01266; DAO; 1.				
DR	PROSITE; PS00626; RCCL2; UNKNOWN.1.				
SQ	SEQUENCE 486 AA; 55220 MM; 5861D01F87F8C830 CRC64;				
Query Match					
Best Local Similarity		16.2%;	Score 376.5;	DB 2;	Length 486;
Matches 129;		Conservative 59;	Mismatches 181;	Indels 75;	Gaps 17;
QY	2	VTKSSLLIVGAGTGTCTALHARRGYTNVTVDYP-----VPSAISAGNDVNVK	53		
Db	3	LNKDSKIIIVGAGVFGLSNALHAKSGYKNVTVDRLDFGANHYTLLEGADTASADINKI	62		
QY	54	ISSGOYSNNKDEIVNEILLAEAFNGWKN-----DPLFKPYHYDHTGL-LMSAC---	100		
Db	63	FRA-QYAEKKHYOE-----LAFRAFEIWEQWDRKIQLLPESEVKEF---TGLRLDLDCGML	114		
QY	101	---SQGLDRLGVVR---VRPGEDPNLVELTRPQFRKLAPGV-----LQGDPPGW	145		
Db	115	RLDDQVGMELASRDNFAREGLGLRFDINDKNDLRAKAQGYGPKVQFALDLOSKIPLD	174		
QY	146	KGYFASGAGWAHARNALVAAREAQRMGVKFTVGTTPQGRVVT-LIFENNVDKAVTGDG	204		
Db	175	EGVF-DSTGMLYASRCLOVAAVLCMSGVRFTMGDGKGTFFKEYVMISNKIVGIVTRDG	233		
QY	205	KIWRABRTFLCAGASAGQFLDFKNQLRPTAWTLVHIALKPEERA-----LYKNIPVFN	258		
Db	234	KKHEAEVLVINCWPSTSLVPELDGINEGTGGNLIKCPDPSKDLKVKYSSKNFPII-G	292		
QY	259	IERG-----FFPEPDEERGBIKICDEHPGYTNMVQSADGTWMSI-----PF	299		
Db	293	WKSCHSREHDYLAGMFIFFVIEPEGILKIIVRTKTNPEKLANGRVVSIPTKTSNPPF	352		
QY	300	EKTQIPKEAETRVALLKETMPOLADRPFSFARICWADTANREFIDRHPQYHSLVLC	359		
Db	353	E--ALPRHIVYQVREWLQTFPDLIDLKWE-SRLWTDTINNDYIYDYVNNKGLFVAC	409		
QY	360	GASGRGFKYLPISIGNLIVDAMEGK	383		
Db	410	GGSGHAFKMLPVLGQFLVDKIEGR	433		
RESULT 15					
Q6BN42	PRELIMINARY; PRT; 446 AA.				
ID	Q6BN42				
AC	Q6BN42				
DT	25-OCT-2004 (T-EMBLrel. 28, Created)				
DT	25-OCT-2004 (T-EMBLrel. 28, Last sequence update)				
DT	25-OCT-2004 (T-EMBLrel. 28, Last annotation update)				
DE	Similar to tr Q43029 Schizosaccharomyces pombe Putative fructosyl				
DE	amino acid oxidase.				
GN	ORFNames=DEHA0F00550g;				
OS	Debaryomyces hansenii CBS767.				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae; Debaryomycetes.				
OX	NCBI_TaxID=284592;				
RN	[1]				
SEQUENCE FROM N.A.					
RP	STRAIN=CBS767;				
RC	Genolevures;				
RG	Genolevures;				
RA	Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,				
RA	Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,				
RA	Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,				
RA	Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,				
RA	Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,				
RA	Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,				
RA	Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,				
RA	Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,				
RA	Nicaud J.M., Nikolaki M., Oztas S., Ozier-Kalogeropoulos O.,				
RA	Pellenn S., Potier S., Richard G.F., Straub M.L., Suleau A.,				
RA	Svenene D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,				
RA	Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,				
RA	Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,				
RA	Wincker P., Souciet J.L.;				
RT	"Genome evolution in yeasts."				
RL	Nature 430:35-44(2004).				
RN	[2]				

```

RA Boissrame A., Boyer J., Cattolico L., Confaniolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennens D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382138; CAG88680.1; -.
DR InterPro; IPR006076; Rad oxred.
DR InterPro; IPR002025; NAD_BS.
DR Pfam; PF01266; DAO; 1.
SQ SEQUENCE 446 AA; 50897 MW; 6C89A7DBBAB75C6B CRC64;

Query Match      15.1%; Score 351; DB 2; Length 446;
Best Local Similarity 26.8%; Pred. No. 5.4e-19;
Matches 126; Conservative 74; Mismatches 186; Indels 84; Gaps 22;

QY      5 SSSLLIVGAGTGTSTALHLARRGYTNVT--VLDPY--PVPSAISAGNDVKNVISSQYS 60
DB      4 TKOIIIVGAGTGLSTALALAE--YANTTHVDRFEPFVPGSSV--DTTRCIR----P 56

QY      61 NKKDEIEVNEILAEAFNGWKNDFLPKPYHDTGL-----LMSACSOEGL 105
DB      57 DYKDETYCR--LAHESMNLKADPMISKLFHQOQMTFVYDGNHDTWEDIFLSSKSSAE-- 112

QY      106 DRLGVRVRPGEDPNLVELTRPEQ-FRKL--APEGVLQGDPPG---W-KGYFARSAGWA 157
DB      113 -----KVNREEGKMDYYDSAEDEVFSIHNRSPNSVETTELGGPKWKNGYTNRVN-GFI 166

QY      158 HARNAL---VAAAREAQRMGVKFTGTGPQGRVVTLIFENNDVKGAVTGDGKIWAERTFL 214
DB      167 DARASIKVYYERARNFPNVNRF-----QVEKIDYFKNTKKVKGVLKNGDKISADLVIV 222

QY      215 CAGASAGQFLDFKNQLRPTAWTLVHIALKPEERALYKNIPIVFNIERGF-FFEPDEERGE 273
DB      223 AAGAWSCKLVNLDNISKSSAIEVVWFKVTPSMEKEWQMSITTNLTSTGINIFPP--YNGE 280

QY      274 IKICDEHPGYTNMQ-----SADGTMSIPEKTOIPKEAETRVRLIKET 319
DB      281 IKVLRSGAGYKNTVAIESPSPFSRVKQEISVPTIIDSP--KOWIPEAEIEINGNLKEI 338

QY      320 MPOLADRPSPFARICWCADTANREFLIDRHPQHSVLVGGCAGRGFKYLPISIGNLIYDA 379
DB      339 MPSELCELPFDRTKLCWLTQTTSNFIIDYPEAENVLVATGSSAHAWKFFIIGDKVVDV 398

QY      380 MEGKVPQKIHELKKNPDIAANRNWRD-----TLGRFGGPNRVMDFDHVK 425
DB      399 IEGRLTELEL--IRKW-----SWKEKLTSTVENENAPRMKGDPKEMKE 438

```

Search completed: May 27, 2005, 13:47:27
Job time : 98.9723 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:35:35 ; Search time 33.5437 Seconds
(without alignments)
972.511 Million cell updates/sec

Title: US-10-622-893A-3
Perfect score: 2329
Sequence: 1 AVTKSSSLIVGAGTGTST.....MDFHDVKEWTVQYRDISKL 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1981.5	85.1	437	2	US-09-031-059-1
2	1981.5	85.1	437	2	US-09-031-059-3
3	160	6.9	433	4	US-09-711-164-412
4	155	6.7	387	3	US-09-457-302-1
5	148.5	6.4	866	4	US-09-949-016-10219
6	146.5	6.3	448	4	US-09-252-991A-21043
7	135.5	5.8	378	3	US-09-199-637A-367
8	135.5	5.8	419	4	US-09-252-991A-21423
9	130.5	5.6	484	4	US-09-252-991A-24053
10	125	5.4	652	4	US-09-252-991A-30872
11	123.5	5.3	476	4	US-09-489-039A-12890
12	123.5	5.3	486	4	US-09-684-405-2
13	116.5	5.0	431	1	US-08-391-339-18
14	116.5	5.0	431	1	US-08-484-274A-18
15	115	4.9	505	4	US-09-902-540-9986
16	114.5	4.9	474	4	US-09-543-681A-6873
17	112.5	4.8	626	4	US-09-331-568A-27
18	110.5	4.7	375	4	US-09-328-352-6891
19	110.5	4.7	635	4	US-09-543-681A-7499
20	108	4.6	380	4	US-09-489-039A-7473
21	108	4.6	430	1	US-08-391-339-5
22	108	4.6	430	1	US-08-484-274A-5
23	106.5	4.6	486	4	US-09-543-681A-7036
24	106	4.6	448	4	US-09-489-039A-7529
25	104.5	4.5	496	3	US-09-147-009-8
26	104.5	4.5	649	4	US-09-252-991A-18996
27	101.5	4.4	438	4	US-09-252-991A-26523

Sequence 27656, A
Sequence 12, Appl
Sequence 80, Appl
Sequence 5943, Ap
Sequence 246, App
Sequence 3, Appli
Sequence 17272, A
Sequence 5, Appli
Sequence 5835, Ap
Sequence 135, App
Sequence 24791, A
Sequence 14124, A
Sequence 18753, A
Sequence 10385, A
Sequence 8135, Ap
Sequence 20386, A
Sequence 16908, A

ALIGNMENTS

RESULT 1
US-09-031-059-1
; Sequence 1, Application US/09031059
; Patent No. 5948659
; GENERAL INFORMATION:
; APPLICANT: KATO, NOBUO
; APPLICANT: SAKAI, YASUYOSHI
; APPLICANT: TANI, YOSHIKI
; APPLICANT: FUKUYA, HIROSHI
; TITLE OF INVENTION: RECOMBINANT FRUCTOSYL AMINO ACID
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,059
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/899,172
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 0020-4253P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-031-059-1

Query Match 85.1%; Score 1981.5; DB 2; Length 437;
Best Local Similarity 82.6%; Pred. No. 1.4e-200;
Matches 360; Conservative 43; Mismatches 32; Indels 1; Gaps 1;

```
QY 2 VTKSSLLIVGAGTGTSTALHLARRGYNVTVDLPYPVPSAISAGNDVKNKVISSQOYSN 61
Db 3 VTKSSLLIIGAGTGTSTALHLARRGYNVTVDLPYPVPSAISAGNDVKNKVISSQOYS 62
QY 62 NKDEIEVNEILAEAFNGWKNDPLFKPYHYDHTGLMSACSQEGLDRLGVRVRPGEDPNLV 121
Db 63 KKDEVEVNEILAEAFNGWKNDPLFKPYHYDHTGLMSACSQEGLDRLGVRVRPGEDPNLV 122
QY 122 ELTRPEQFRLAPEGVQLQDFFGKGYFARSAGAGWAHARNALVAAAREAQRMGVKFTGT 181
Db 123 ELTRPEQFRLAP-GVLEKGFPGWRGYHRSNAGWAHARNALVAAAREAQRMGVKFTGT 181
QY 182 POGRRVVTLLIFENNDVKGAVTGDKIWRABEQTILCAGAAAGQFLDFKQRLRPTAHTLVHIA 241
Db 182 POGRRVVTLLIFENNDVKGAVTGDKIWRABEQTILCAGAAAGQFLDFKQRLRPTAHTLVHIA 241
QY 242 LKPEERALKYNIPIFNIERGFFFEDEERGEIKICDEHPGYTNMVQSDGTMSPFEK 301
Db 242 LKPEERALKYNIPIFNIERGFFFEDEERGEIKICDEHPGYTNMVQSDGTMSPFEK 301
QY 302 TQIPKEAETRVALLKETWPLQADRPFPARICWCADTANREFLIIDRHPQYHSLVLGCGA 361
Db 302 TQVPREAENVRKLLSETWPLQADRPFPARICWCADTANREFLIIDRHPQYHSLVLGCGA 361
QY 362 SGRGFKYLPISIGNLIVDAMEGKVPQKIHELKWNPDIAANRWDRDTLGRFGGPNRVMDFH 421
Db 362 SGRGFKYLPISIGNLIVDAMEGKVPQKIHELKWNPDIAANRWDRDTLGRFGGPNRVMDFH 421
QY 422 DVKEWNTVQYRDISKL 437
Db 422 EVKEWNTVQYRDISKL 437
```

RESULT 2

```
US-09-031-059-3
; Sequence 3, Application US/09031059
; Patent No. 5948659
; GENERAL INFORMATION:
; APPLICANT: KATO, NOBUO
; APPLICANT: SAKAI, YASUYOSHI
; APPLICANT: TANI, YOSHIKI
; APPLICANT: FUKUYA, HIROSHI
; TITLE OF INVENTION: RECOMBINANT FRUCTOSYL AMINO ACID
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,059
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/899,172
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 0020-4253P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-059-3

Query Match      85.1%; Score 1981.5; DB 2; Length 437;
Best Local Similarity 82.6%; Pred. No. 1.4e-200;
Matches 360; Conservative 43; Mismatches 32; Indels 1; Gaps 1;

QY 2 VTKSSLLIVGAGTGTSTALHLARRGYNVTVDLPYPVPSAISAGNDVKNKVISSQOYSN 61
Db 3 VTKSSLLIIGAGTGTSTALHLARRGYNVTVDLPYPVPSAISAGNDVKNKVISSQOYS 62
QY 62 NKDEIEVNEILAEAFNGWKNDPLFKPYHYDHTGLMSACSQEGLDRLGVRVRPGEDPNLV 121
Db 63 KKDEVEVNEILAEAFNGWKNDPLFKPYHYDHTGLMSACSQEGLDRLGVRVRPGEDPNLV 122
QY 122 ELTRPEQFRLAPEGVQLQDFFGKGYFARSAGAGWAHARNALVAAAREAQRMGVKFTGT 181
Db 123 ELTRPEQFRLAP-GVLEKGFPGWRGYHRSNAGWAHARNALVAAAREAQRMGVKFTGT 181
QY 182 POGRRVVTLLIFENNDVKGAVTGDKIWRABEQTILCAGAAAGQFLDFKQRLRPTAHTLVHIA 241
Db 182 POGRRVVTLLIFENNDVKGAVTGDKIWRABEQTILCAGAAAGQFLDFKQRLRPTAHTLVHIA 241
QY 242 LKPEERALKYNIPIFNIERGFFFEDEERGEIKICDEHPGYTNMVQSDGTMSPFEK 301
Db 242 LKPEERALKYNIPIFNIERGFFFEDEERGEIKICDEHPGYTNMVQSDGTMSPFEK 301
QY 302 TQIPKEAETRVALLKETWPLQADRPFPARICWCADTANREFLIIDRHPQYHSLVLGCGA 361
Db 302 TQVPREAENVRKLLSETWPLQADRPFPARICWCADTANREFLIIDRHPQYHSLVLGCGA 361
QY 362 SGRGFKYLPISIGNLIVDAMEGKVPQKIHELKWNPDIAANRWDRDTLGRFGGPNRVMDFH 421
Db 362 SGRGFKYLPISIGNLIVDAMEGKVPQKIHELKWNPDIAANRWDRDTLGRFGGPNRVMDFH 421
QY 422 DVKEWNTVQYRDISKL 437
Db 422 EVKEWNTVQYRDISKL 437
```

RESULT 3

```
US-09-711-164-412
; Sequence 412, Application US/09711164
; Patent No. 6589738
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERAPY
; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711,164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 412
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-711-164-412

Query Match      6.9%; Score 160; DB 4; Length 433;
Best Local Similarity 23.3%; Pred. No. 6.6e-08;
Matches 106; Conservative 49; Mismatches 146; Indels 154; Gaps 24;

QY 8 LLIVGAGTGTSTALHLARRGYNVTVDLPYPVPSAISAGNDVKNKVISSQOYSNNKDEIE 67
Db 17 ILLIIGAGTGTSTALHLARRGYNVTVDLPYPVPSAISAGNDVKNKVISSQOYSNNKDEIE 67
QY 68 VNEILAEAFNGWKNDPLFKPYHYDHTGLMSACSQEGLDRL---GVRVRPGEDPNLV 124
```



```
; Sequence 21043, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21043
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21043

Query Match          6.3%; Score 146.5; DB 4; Length 448;
Best Local Similarity 23.7%; Pred. No. 1.9e-06;
Matches 98; Conservative 59; Mismatches 164; Indels 93; Gaps 23;

QY 6 SSSLLIVGAGTGTSTALHLARRGYTNVTLDPPVPSPSAISAGNDVNKVISSGOYNNKDE 65
DB 51 TDVCVIGAGTGTSTALFLENGF-KVTVLETAQVGFAS-GRNGGOIVNS--YSRDIDV 106
QY 66 IEVN-----ETLAERAFNGWKNDPLFKPYHYHTDGL-----LMSACSQEGLDLRLGVR 111
DB 107 IERTVGKREAQLGENAFEGGR---IIRBRVARYIGCOLKGGVFAAFTEKQMDHLRAQ 163
QY 112 VRGE--DPNLVELTRPEQFKLAPGVLQDPPGKVGYPARSAGAWAHARNAVAAARE 169
DB 164 KQWERYGHNLIMDAKRIR-----EVVATDNYIG--GMLDMSG-GHIHPLNALGEAAA 216
QY 170 AQRMGVKFTGTGQGRVVTLLIFENNNDVKGAVTGDGKIKWRAERTFLCAGASAGOLFDFKNQ 229
DB 217 VESLGGRIYEQSPATRI-----ERGASPVVHTPGKV---KAKFIVVAGNA--YLN--- 262
QY 230 LRPTATWLVIHAKPERALYKNIP-----VI-----FNIERGFFPEDEERGEIKICDEHP 281
DB 263 -----GLVPELAA--KSMPCGTQVITTEPLNEELAHSLLPQD-----YCVEDC 303
QY 282 GYTNMVQSGDGTWMSI-----PPEKTOIPKEAETRVALLKETMPQLADRPFSFARICWCA 337
DB 304 NYLLDYRLSGDKRLRIGYGVIGADGPDADIEAIRPKMLKTPPQDKDKIDFA---W--- 358
QY 338 DTANRFLIDRHPQ-----YHSLVLGCGASGRGFKYLPSTIGNLIVDAMEGK 383
DB 359 -TCGNFLLTSLRLOVQGRIGNDIYYSQ-----GCSGHGVYTYHVAGKVLAEALRQ 407

RESULT 7
US-09-199-637A-367
; Sequence 367, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsengalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
```

```
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 367
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-367
```

```
Query Match          5.8%; Score 135.5; DB 3; Length 378;
Best Local Similarity 24.3%; Pred. No. 2.1e-05;
Matches 102; Conservative 45; Mismatches 165; Indels 107; Gaps 19;

QY 5 SSSLLIVGAGTGTSTALHLARRGYTNVTLDPPVPSPSAISAGNDVNKVISSGOYNNK 64
DB 16 SRDVVVGAGVIGLLTARELALAGRLVTLVERGESGREASWAGGGIVSPLYPWRYS 72
QY 65 EIEVNEILAEAFNGWKNDPLFKPYHYHTDGLLMSACSQEGLDLRLGVRVRPGE 120
DB 73 -----AVTALAHWSQD--FYP-----ALGQRLLDDETGL-----DPEVHTVG 106
QY 121 ---VEL-----TRPEQFKLAPGVLQDPPGKVGYPARS--GAGWAHARNA 162
DB 107 LWLDDDDTQTEALQWNRNTRP---LKEVPIEBAVAAVFLGAGFORAVTMSGVANRNP 163
QY 163 LVAAAREAQRMGVKFTGTGQGRVVTLLIFENNNDVKGAVTGDGKIKWRAERTFLCAGASAG 222
DB 164 RLARSILRASLQFANLELHEQTEVGRWLRDGRVVGVAISRGEI-RGDKVLLAAGAWSGE 222
QY 223 FLDFKNQLRPTATWLVIHAKPERALYK---NIPVIFNIERGFFPEDEERGEIKICD 278
DB 223 L-----LKPLGLIELPVPVVK-QOMILYKCAADFLPRMV-LAKGRYAIPRRD-GHILIGS 273
QY 279 --EHPGYTNMVQSGDGTWMSIPEKTOIPKEAETRVALLKETMPOLAD-----R 326
DB 274 TLEHSG-----FDKTPDTEALES-LRASAELLPELADMQPVVAHWAGLR 316
QY 327 PFSFARICWADTANRFLIDRHPQYHSLVLGCGASGRGFKYLPSTIGNLIVDAMEGKVP 385
DB 317 PGSPEGIPY-----IGVPVPGDGLWLTNGHVRNGLVLAPASCRLLADLMSGREP 365
```

```
RESULT 8
US-09-252-991A-21423
; Sequence 21423, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21423
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21423
```

```
Query Match          5.8%; Score 135.5; DB 4; Length 419;
Best Local Similarity 24.3%; Pred. No. 2.4e-05;
Matches 102; Conservative 45; Mismatches 165; Indels 107; Gaps 19;

QY 5 SSSLLIVGAGTGTSTALHLARRGYTNVTLDPPVPSPSAISAGNDVNKVISSGOYNNK 64
DB 57 SRDVVVGAGVIGLLTARELALAGRLVTLVERGESGREASWAGGGIVSPLYPWRYS 113
QY 65 EIEVNEILAEAFNGWKNDPLFKPYHYHTDGLLMSACSQEGLDLRLGVRVRPGE 120
```


Db 114 -----AVTALAHSQD--FYP-----ALGQRLDETGL-----DPEVTVG 147
Qy 121 ---VEL-----TRPEQFRKLAPGVLOQDPGWMKGFYARS--GAGWAHARNA 162
Db 148 LYWLDDLDQDEALQWARKHTRP---LKEVPTEEAAYAAPVGLGAGFORAVYMSGVANVRNP 204
Qy 163 LVAABAREAQRMGVFTVTGTPQGRVVTLIFENNVDKGVATGDKIWAERTFLCAGASAGQ 222
Db 205 RLASRLRASLQQFANLELHEQTEVRGMLRDGDRVLGVATSRGEI-RGDKVLLAAGAWSGE 263
Qy 223 FLDFKNRLPTAWTLVHIALKPEERALKYK----NIPVFNIERGFFPEPEERKEICD 278
Db 264 L-----LKLGLGLPVPVK-QQWILYKCAADFLPMV-LAKGRYAIPRRD-CHILIGS 314
Qy 279 --EHPGVNMQSADGTMSTIPPEKTOIPKEATRVALLKETMPQLAD-----R 326
Db 315 TLEHSG-----FDKTPDDEALS-LRASAAELLPELADMQPVAAHWAGLR 357
Qy 327 PFSFARICWCACTANREFLIDRHPQYHSLVLCGASGRGPKYLSIGNLIVDAMEGKVP 385
Db 358 PGSPEGIPY-----IGPVFGDGLMNTGHRNGLVLAPASCRLLADLMSGREP 406

RESULT 9
US-09-252-991A-24053
; Sequence 24053, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24053
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24053

Query Match 5.6%; Score 130.5; DB 4; Length 484;
Best Local Similarity 22.2%; Pred. No. 0.0001;
Matches 101; Conservative 70; Mismatches 183; Indels 101; Gaps 23;

Qy 4 KSSLLIVGAGTGTSTALHLAGRGYTNVTVDLPYPVPSAIS-----AGNDVYNK- 52
Db 81 EQADVCVGGFGSVNTALELAQRGP-SVLLAHRIWGASGRNGGQLIRGVGHVDEQF 139
Qy 53 --VTSSQYNNKDEIEVNEILAEAFNGWKNDPLFPYHYDTGLMSACSQEGDLRLGV 110
Db 140 FVPVGA-----DGVKALKMLGLEAVEIVRR--RVEQYAIIDCLRWGYCD-----L 182
Qy 111 RVRPG-----EDPNLVELTRPEQFRKLAP-----EGVLOQDFPGWKGYFARSAGWAHAR 160
Db 183 ANKFGDYQGFREDMEELQALGYRHEMLVPAENRSVVGSD--RYVGLVDMGSGHLHPL 240
Qy 161 NALVAAAREAQRMGVKFTVTGTPQGRVVTLIFENNVDKGVATGDKIWAERTFLCAGASA 220
Db 241 NLVLGEAAAQSLGVLRFERSP-----VTRIDYGAEVQ-VHTATGKV-RAKT--LVLCNA 292
Qy 221 GQFLDFKNQLRP-----TAWTLVHIALKP-ERALKYKNIPVFNIERGFFPEPEDEERG 272
Db 293 -----YNDNLNPLGGKVLPAQSYV-IATFPLDQLARQLLP-----Q 329
Qy 273 EIKICDHPGVNMQSADGTM--SIPPEKTOIPKEATRVALLKETMPQLADRPFSF 330
Db 330 NMAVCDORVALDYRLSADNRLLPGGACHYSGRDPFSDIAAYMRPKMLEVFPQLANVRIDY 389

Qy 331 ARICWCACTANREFLIDRHPQYHSLVLCGASGRGPKYLSIGNLIVDAMEG----- 382
Db 390 QWGMIGIGANRLPQIGRLFCQPNVYFAQYSGHGVNATHLAGQLLAEAIIGSQSDGFDL 449
Qy 383 --KVPOKIH---ELIKWNPDIANRNW---RDTLG 409
Db 450 FAKVPHITFFGGKLLR-SPILLALGMWYRLKEKLG 483

RESULT 10
US-09-252-991A-30872
; Sequence 30872, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30872
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30872

Query Match 5.4%; Score 125; DB 4; Length 652;
Best Local Similarity 21.8%; Pred. No. 0.00065;
Matches 92; Conservative 54; Mismatches 162; Indels 114; Gaps 16;

Qy 5 SSSLLIVGAGTGTSTALHLAGRGYTNVTVDLPYPVPSAISAGNDYNNKISSQYNNK 63
Db 283 NAEVIVVGAGIVGSACAHELARRG-LDVLVDSRRGGATAVGMGHLV-----AMDDNP 334
Qy 64 DEIEVNEILAEAFNGWKNDPLFPYHYDTGLMSAC-----SQEGDLRLGVVR 113
Db 335 AELALSD-YSIQAWRTWAADLPEDCAYRNCGTWLAADAELAEAEKRCQALLAAGV--- 390
Qy 114 PGSDPNLVELTRPEQFRKLAPGVLOQDFPGWKGYFARSAGWAHARNALVAAAREAQRM 173
Db 391 -----ACEMLDAARLRLDLEP--VLR--FGLAGALKVPDGDILYAPN---AARWLLERA 436
Qy 174 GVXFTVTGTPQGRVVTLIFENNVDKGA--VTGDKGIWAERTFLCAGASAGQFLDFKNQLR 231
Db 437 G-----PRLRLHAEVSEVDGSRRLADGRWLSAEALVLANGTHAGEL----- 479
Qy 232 PTAWTLVHIALKPEERALKYKNIPVFNIERGFFPEPEERKEIKICDEHP----- 281
Db 480 -----CAELIRP-----KKGHLITDYPGTLRHQLVEL 509
Qy 282 GYTNMQSADGTM-----SIPPEKTOIPKEATRVALLK--ETMPQL 323
Db 510 GYVSSAHASGTSVAFNAQPRPTGQVFLGSSRFDTLDPQVGPVLAERMLRRALDYLPL 569
Qy 324 ADPRFSPARICWCACTANREFLIDRHPQYHSLVLCGASGRGPKYLSIGNLIVDAMEGK 383
Db 570 AGLNATRAMTGFRAATPDGLPLGHEHPAQPGMLAVGHEGLGVTTAPGSRALLAAQLFGE 629
Qy 384 VP 385
Db 630 TP 631

RESULT 11
US-09-489-039A-12890
; Sequence 12890, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 12890
LENGTH: 476
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12890

Query Match 5.3%; Score 123.5; DB 4; Length 476;
Best Local Similarity 22.5%; Pred. No. 0.00056;

Matches 92; Conservative 48; Mismatches 148; Indels 121; Gaps 21;

QY 6 SLLIIVGAGTGTSTALHARRGYTNVTLDPY-----PVPISAISAGNDVNKV 53
DB 73 ADVVIIGGFSGINTALEAEQITNVVLEARHLGCGTGRNGGQVMAGI--GHDIEAV 130
QY 54 ISSQGYNNKDEIEVNEILAEAFNGKNDPLFKPYHDTGLLMSACSQEGLD 106
DB 131 -----KKHVGKGLTFLKIANLGAGIIRIRIKYNIDADFVPGY---GYL--AYNQRLK 181
QY 107 RLG-----VRVPGEDPNLVELTRPEQFKLAPEGVLQDFPGWKGYFARSGAGWAHAR 160
DB 182 TLRQWKEKFAATPDEE---IELYQKEVQVGVSEV-----YCGALKHMGGGQIHSL 231
QY 161 NALVAAREAQRMGVKFTGT-----QGRVVTLLIFENNNDVKAVTGDGKIWAERT 212
DB 232 NMLLSAQAAHSLGVKIFESSPVVYNGKQVRVTAM-----GSVKAALLWACD-- 282
QY 213 FLCAGASAGFLDPKQNLRTAM--TLV-----HIALKPEERALKYKNIPVFNIERGFFFE 266
DB 283 -----SFLNNMEPEIYNKTLVTYSQVSTPLSDELIERISPL---RGAFSD 326
QY 267 PDEERGEIKCDHPGYTNMVQASDGTMSIPEKTIQIPKEATRVALLKETMPOLADR 326
DB 327 IRPVINYVTR-----NRLFGSATRF---VEYT--PNDFAAWNRTLLAEVFPYLRDV 376
QY 327 PFSFARICWADTANREFLIDRHPQVHSLVLGCGASGRGFKYLPISGNL 375
DB 377 KIDFA---WGGPNA-----CSAN-----LFPQIGTL 399

RESULT 12
US-09-684-405-2
Sequence 2, Application US/09684405
Patent No. 6770477
GENERAL INFORMATION:
APPLICANT: Dennis J. Slamon
APPLICANT: Juliana J. Oh
TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED GENES
TITLE OF INVENTION: ASSOCIATED WITH HER-2/NEU OVEREXPRESSION
FILE REFERENCE: 30448.79USU1
CURRENT APPLICATION NUMBER: US/09/684,405
CURRENT FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/157,923
PRIOR FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 486
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-684-405-2

Query Match 5.3%; Score 123.5; DB 4; Length 486;
Best Local Similarity 20.9%; Pred. No. 0.00058;
Matches 94; Conservative 53; Mismatches 179; Indels 123; Gaps 20;

QY 6 SLLIIVGAGTGTSTALHARRGYTNVTLDPYVPVPSAISAGNDVNKV-----53
DB 64 SDVVIUGGVLGLSVAYWLKLESRRGAIRVLVVERDHTYSQASTGLSVGGICQFSLPE 123
QY 54 -----ISSQGYNNKDEIEVNEILAEAFNGKNDPLFKPYHDTGLLMSACSQEGLD-R 107
DB 124 NIQLSLFSAFLRN-----INEYLAV-----VDAPPLDLRFNPSGYLLILASEKDAAME 172
QY 108 LGVRYRPGEDPNLVELTRPEQFKLAP-----RGVLOGDFPGWKGYFARSGAGWAHARNAL 163
DB 173 SNVKVQROEGAK-VSLMSPDLRNKFPWINTGVALAS-----YGMEDGWFDPWCLL 224
QY 164 VAAAREAQRMGVKFTGT-----TPQGRVVTL-----IFENNNDVKG 198
DB 225 QGLRRKVQSLGLVFCQGEVTRFVSSSQRMILTDDKAVLVKRIHEVHVKMDRSLEYQVPEC 284
QY 199 A-VTDGKIWAERTFLCAGASAGFLDPKQNLRTAMTLVHIALKPEERALKYKNIPVIF 257
DB 285 AIVINAAGAWSAQIAAL-AGVGEPP-----PGLQGTKLPEVPRKRYVY-----327
QY 258 NIERGFFPEPDEERGEIKICDEHPGY-TNMVQASDGTM-----MSIPFEKTIQIPKEATR 311
DB 328 -----VWHCPQPGLETPLVADTSGAYFRREGLSNYLGGRSPTQECP 371
QY 312 VRALLK-----ETMPOLADRPFSFARI-----CWCA-----DTANREFLIDRHPQVHSL 355
DB 372 DPANLEVDHDFQDKVWPHALARVPAPFTLKVQSAWAGYYDNTFDQGVGPHPLVVNM 431
QY 356 VLGCASGRGFKYLPISGNLIVD-AMEGK 383
DB 432 YFATPGSHGLQQAOPGIGRAVAEMVLKGR 460

RESULT 13

US-08-391-339-18
Sequence 18, Application US/08391339
Patent No. 5463175
GENERAL INFORMATION:
APPLICANT: Kishore, Ganesh M.
APPLICANT: Barry, Gerard F.
TITLE OF INVENTION: Glyphosate Tolerant Plants
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr.,
ADDRESSEE: Monsanto Co. BB4F,
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/391,339
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/156,968
FILING DATE:
APPLICATION NUMBER: US/07/717,370
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner, Dennis R., Jr.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10533)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:

```
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-391-339-18

Query Match      5.0%  Score 116.5; DB 1; Length 431;
Best Local Similarity 22.7%; Pred. No. 0.0026;
Matches 94; Conservative 36; Mismatches 119; Indels 165; Gaps 22;

QY 10 IVGAGTGTSTALHLARRGYNVTVDLPYPVPSAISAGNDVNKVISSGQYNNKDEIEVN 69
DB 10 IAGAGIVGCTALMLQRRGF-KVTLDNPPGEGASFGN-----47
QY 70 EILAEAFNGWKNDPLPKPYHTDTGLMSACSQEGDLRLG-VRVPGEDP-----118
DB 48 ----AGCFNGSSVVPMSMP-----GNLTSV-PKWLLDPMGPLSIRFGYFPTIMPWLIRFL 97
QY 119 -----NLVELTRPEQFRKLAPEGVL-----OGDFPGWKGY--FARSGA 154
DB 98 LAGRPNKVKEQAKALRNLIKSTVP-LIKSLAEADASHLIRHEGLHTVYRGEADFARDRG 156
QY 155 GWA-----HARNALVAAAREAQ-----MGVKFVTG-----TPQGRVVTLIF---191
DB 157 GWELRLNGVRTQILSA--DALRDFDNLSHAFTKGLIEENGHTINPQG-LVTLLPRRF 213
QY 192 -----ENNDVGAVTGDGKIWRAERTFLCAGASQGLDFKNQLRPTA 234
DB 214 IANGGEFVSARVGFETEGRALKGITTTNG-VLAVDAAVVAAGAH-----258
QY 235 WTLVHIALKEERALKNIPIVNIERGFPEDEERG-EIKICDEHPGYTNMVQSADGT 293
DB 259 -----KSLANSLGDDIPL-----DTERGYHIVIANPEAAPRIPTTDASGK 298
QY 294 MMSIPFEK-----TQIP--KEAE--TRVRALKETMPOLADRPFS 329
DB 299 FIATPEMGLRVAGTVEFAGLTAAPNWKRAHVLYTRARKLLPALAPASSEERYS 352

RESULT 14
US-08-484-274A-18
; Sequence 18, Application US/08484274A
; Patent No. 576760
; GENERAL INFORMATION:
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Glyphosate Tolerant Plants
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,274A
; FILING DATE: 07 June 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Patterson, Melinda L.
; REGISTRATION NUMBER: 33,062
; REFERENCE/DOCKET NUMBER: MOBT130 38-21(13560)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713)789-2679
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-274A-18

Query Match      5.0%  Score 116.5; DB 1; Length 431;
Best Local Similarity 22.7%; Pred. No. 0.0026;
Matches 94; Conservative 36; Mismatches 119; Indels 165; Gaps 22;

QY 10 IVGAGTGTSTALHLARRGYNVTVDLPYPVPSAISAGNDVNKVISSGQYNNKDEIEVN 69
DB 10 IAGAGIVGCTALMLQRRGF-KVTLDNPPGEGASFGN-----47
QY 70 EILAEAFNGWKNDPLPKPYHTDTGLMSACSQEGDLRLG-VRVPGEDP-----118
DB 48 ----AGCFNGSSVVPMSMP-----GNLTSV-PKWLLDPMGPLSIRFGYFPTIMPWLIRFL 97
QY 119 -----NLVELTRPEQFRKLAPEGVL-----OGDFPGWKGY--FARSGA 154
DB 98 LAGRPNKVKEQAKALRNLIKSTVP-LIKSLAEADASHLIRHEGLHTVYRGEADFARDRG 156
QY 155 GWA-----HARNALVAAAREAQ-----MGVKFVTG-----TPQGRVVTLIF---191
DB 157 GWELRLNGVRTQILSA--DALRDFDNLSHAFTKGLIEENGHTINPQG-LVTLLPRRF 213
QY 192 -----ENNDVGAVTGDGKIWRAERTFLCAGASQGLDFKNQLRPTA 234
DB 214 IANGGEFVSARVGFETEGRALKGITTTNG-VLAVDAAVVAAGAH-----258
QY 235 WTLVHIALKEERALKNIPIVNIERGFPEDEERG-EIKICDEHPGYTNMVQSADGT 293
DB 259 -----KSLANSLGDDIPL-----DTERGYHIVIANPEAAPRIPTTDASGK 298
QY 294 MMSIPFEK-----TQIP--KEAE--TRVRALKETMPOLADRPFS 329
DB 299 FIATPEMGLRVAGTVEFAGLTAAPNWKRAHVLYTRARKLLPALAPASSEERYS 352

RESULT 15
US-09-902-540-9986
; Sequence 9986, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9986
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-9986

Query Match      4.9%  Score 115; DB 4; Length 505;
Best Local Similarity 21.9%; Pred. No. 0.0049;
Matches 92; Conservative 54; Mismatches 148; Indels 126; Gaps 24;

QY 2 VTKSSLLIVGAGTGTSTALHLARRGYNVTVDLPYPVPSAISAGNDVNKVISSGQYS- 60
DB 115 VPAKAKVVIIGGIMGLALAYNLARAGETDVVVLERYLCAGASGRNGGVGRVWQGTPL 174
QY 61 --NNKDEIEVNEILABE-AFNGW-----KNDFLPKPYHYDHTGLMSACSQEGDLR 107
DB 175 VELAKESIEMKGFARELGINVWLROGGYIFLAKTAPVAQRLERNVSL-----HNR 225
QY 108 LGVRVRPGEDPNLVELTRPEQFRKLAP-----EGVLQDFPGWKGYFARS 152
```


QY 0Z NKDEIEVNEI DAEERAFN

Db 423 MDLPKSDVEGWNIK 437

RESULT 8

ABG30783

ID ABG30783 standard; protein; 440 AA.

XX

AC ABG30783;

XX

DT 05-NOV-2002 (first entry)

XX

DE Fusarium oxysporum polypeptide #2.

XX

KW Glycoamino acid; glycoprotein; deoxycholic acid; deoxycholic amide;

KW cholic acid amide octyl glucoside; quaternary ammonium salt;

KW quaternary ammonium salt cationic surfactant; concanavalin A; betaine;

KW ascorbic oxidase; 4-(2-hydroxyethyl)-1-piperazinyl group; albumin;

KW protease; glycoalbumin; bromocresol purple; glycation; diabetes;

KW ascorbic acid; mutant; muten.

XX

OS Fusarium oxysporum.

OS Synthetic.

XX

XX

XX

FT Key Location/Qualifiers

FT Misc-difference 246 /note= "Encoded by ATG"

FT Misc-difference 372 /note= "Wild-type Lys substituted by Trp. Encoded by TTG"

FT Misc-difference 396 /note= "Encoded by TTC"

FT

XX WO200261119-A1.

XX

XX 08-AUG-2002.

XX

PF 30-JAN-2002; 2002WO-JP000721.

XX

XX 31-JAN-2001; 2001JP-00022953.

PR 16-FEB-2001; 2001JP-00039796.

PR 08-AUG-2001; 2001JP-00240002.

XX

PA (ASAH) ASahi KASEI KOGYO KK.

XX

PI Kouzuma T, Yoshioka I, Arai M, Sumitani J, Imamura S;

XX

XX WPI: 2002-599854/64.

DR N-PSDB; ABK30476.

DR

XX Compositions for assaying glycoprotein without interference from globulin and ascorbic acid in blood, useful in clinical examination including diagnosis and management of diabetes.

PT

XX Disclosure; Page 66-69; 82pp; Japanese.

XX

CC The invention relates to compositions for assaying glycoprotein with use of protease and enzymes acting on glycoamino acids in the presence of at least 1 of deoxycholic acid or amide, cholic acid amide octyl glucoside, quaternary ammonium salt, quaternary ammonium salt cationic surfactant, concanavalin A and betaine and/or ascorbic oxidase and a buffer without a 4-(2-hydroxyethyl)-1-piperazinyl group. The invention also relates to compositions for assaying albumin containing the protease and enzymes for determining glycoalbumin and a protein-denaturing agent and/or a compound having S-S bond and bromocresol purple for evaluation of glycation proportion of albumin into glycoalbumin with respect to the other albumins measured separately. The compositions are useful in clinical examination including diagnosis and management of diabetes. The compositions can be used in methods which avoid the effects of globulin and ascorbic acid components and stabilise enzymes such as protease acting on glycoamino acids. This sequence represents a polypeptide used in the scope of the invention

XX

SQ Sequence 440 AA;

Query Match 49.3%; Score 1148.5; DB 5; Length 440;

Best Local Similarity 51.0%; Pred. No. 4.1e-107;

Matches 222; Conservative 75; Mismatches 131; Indels 7; Gaps 5;

QY 2 VTKSSLLIVGAGTGTSTALHLARRGYTNVTLDPYPVPSAISAGNDVNVKVIS--SGQY 59

Db 4 LTKSQILLIVGGTGWGCTSTALHLARRGYTNVTLVDNVRIPSPISAGHDVNVKLAGRLSTAD 63

QY 60 SNNKDEIEVNEILABEAFNGWKNDFLKPYPYHDTGLMSACSQEGLDRLGVRVRPGED-P 118

Db 64 SKGDDSDSIWKALSYAAAGQWLHDFVFPQPCHTGSSVAGSTPKSIKQL-VEDEIGDDID 122

QY 119 NULVELTRPEQFKLAPEGVLOQDFPGWKGYFARSAGAHARNALVAAAREAORMGVKVF 178

Db 123 QYTPLNTAEDFRKTPMEGILTCNFPWKGYFARSGWVHARKAMKAAPFEESERLGVKFI 182

QY 179 TCTPGRVVTLFENNVDYKGVATGDKIWAERTFLCAGASAGQFLDKNQLRPTAWTILV 238

Db 183 TGSPGKVESLIPEDGDVRGAKTAGDKEHRAKRTILSAGASAEFFLDFENQIQPTAWTILG 242

QY 239 HIALKPEERALYKNIPVFNIERGFFFEDEPDERGEIKICDEHPGYTNMVQSADGTW-MSI 297

Db 243 HIQITPEETKLYKNLPPLFNINQGFMEPEDHLQKMCDEHPGYCNWVEKPSKYQPSI 302

QY 298 PFEKTIQPKAEATRVALLKETMPOLADRPFPFARIWCADTANREFLIDRHPQYHSLVL 357

Db 303 PFAKHQVPTAEARRMKQFLKDIMPOLADRPVHARICWCADTQDRMFLITYHPRHPSLVI 362

QY 358 GCGASGRGPKYLPISGNLIVDAMEGKVPQKIHILIKWNPDIANRNWRDTLGRFGGPNRV 417

Db 363 ASGDCGTGYWHITSIGKFI SDCWEGTEERPAKYWRWRPEKTEFWGKDLDRFGADDKI 422

QY 418 MDF--HDVKEWTNVQ 430

Db 423 MDLPKSDVEGWNIK 437

RESULT 9

ADM78789

ID ADM78789 standard; protein; 444 AA.

XX

AC ADM78789;

XX

DT 01-JUL-2004 (first entry)

XX

DE Fusarium proliferatum fructosylamine oxidase Q2 protein sequence.

XX

KW fructosylamine oxidase; fructosyllysine; fructosylvaline;

KW Amadori compound; medical examination; diagnosis; serum glucose level;

KW food inspection; enzyme.

XX

OS Fusarium proliferatum.

XX

XX WO2004029251-A1.

PN

XX

PD 08-APR-2004.

XX

XX 16-SEP-2003; 2003WO-JP011766.

XX

PR 24-SEP-2002; 2002JP-00277214.

PR 24-OCT-2002; 2002JP-00309734.

XX

XX (ARKR-) ARKRAY INC.

PA

XX Yoshida N, Tani Y, Yonehara S;

PI

XX WPI: 2004-316127/29.

DR N-PSDB; ADM78788.

DR

XX Fusarium proliferatum-originated fructosylamine oxidase useful in measurement of Amadori compound and particularly applicable in medical examination including diagnosis and management of diabetes.

PT

XX

QY	225	DFXQLRPTAWTLVHIALKPEERALKYNIPIVFNIERGFFPEDEBERGEIKICDEHPGYT	284
Db	229	DLEDQCVSKAWVFAHIQLTPQESAQKVDVVDYDGYGFFFEFN-EHGVLKVCDEFFGFS	287
QY	285	NM-----VQSADGTWMSIPEKTO-----IPKEAETRVALLKETMPQLADRFPSFARIC	334
Db	288	RFXLHPQPYGATSKLISVPSRSHAKHPTDTYDPSSEETIRKAIARFPRFKDKELFNRSMC	347
QY	335	WCADTANREFLIDRHPQYHSLVLCGASGRGFKYLPISIGNLIYDAMEGKVPQKIHELIIKW	394
Db	348	WCTDTADANLLICEHPKWNFILATGDSGHSFKVLENIQKHVVLELLEGRLPQDLAGARW	407
QY	395	NP-----DIAANRWDRDTLGRFGGPNRVMDFHVK	424
Db	408	RPGGDALKSRSAAPAKDLAEMPGWK-----HDAK	436
RESULT 15			
ABR41990			
ID	ABR41990	standard; protein; 437 AA.	
AC	ABR41990;		
XX			
DT	11-AUG-2003	(first entry)	
DE	Eupenicillium terrenum	fructosyl peptide oxidase.	
DE			
XX			
KW	Fructosyl peptide oxidase; enzyme; diabetes; diagnosis.		
XX			
OS	Eupenicillium terrenum.		
XX			
PN	EP1291416-A1.		
XX			
PD	12-MAR-2003.		
XX			
PF	03-SEP-2002; 2002EP-00256108.		
XX			
PR	04-SEP-2001; 2001JP-00266665.		
PR	12-DEC-2001; 2001JP-00378151.		
PR	06-AUG-2002; 2002JP-00228727.		
XX			
PA	(KIKK) KIKKOMAN CORP.		
XX			
PI	Kurosawa K, Hirokawa K, Kajiyama N;		
XX			
DR	WPI; 2003-395517/38.		
DR	N-PSDB; ACC48873.		
XX			
XX	New fructosyl peptide oxidases useful as enzymes for clinical diagnosis.		
PT			
PS	Claim 20; Page 31-34; 59pp; English.		
XX			
CC	The present sequence is the protein sequence of a novel fructosyl peptide oxidase (FPO) of Eupenicillium terrenum ATCC18547. FPO enzymes of the invention act on fructosyl valyl histidine in the presence of oxygen and catalyze a reaction that produces alpha-ketoaldehyde, valyl histidine and hydrogen peroxide. The E. terrenum FPO exhibits 182% relative activity for fructosyl glycine and 9.78% for epsilon fructosyl lysine when compared to 100% activity for fructosyl valyl histidine. It shows optimal activity in the pH range 6.0-8.0 (preferably pH 7.0) and at temperatures around 35 degrees C, and is stable in the pH range 5.0-9.0 and at temperatures up to about 45 degrees C. The Km value for fructosyl valyl histidine is 4.25 mM. The E. terrenum cDNA can be used for mass production of the enzyme by recombinant methods. Stable FPO enzymes such as E. terrenum FPO are useful in clinical diagnosis, e.g. of diabetes		
XX			
XX	Sequence 437 AA;		

Query Match 26.8%; Score 625; DB 6; Length 437;
Best Local Similarity 34.2%; Pred. No. 5.7e-54;
Matches 153; Conservative 81; Mismatches 174; Indels 40; Gaps 14;

QY	1	AVTKSSLLIV--GAGTGTCTSTALHILARRGYT--NVTVLDPYPVPSAISAGNDVNKVISS	56
Db	2	AHSRASTKVVWVGGGGTIGSTALHILIRSGYTPSNITVLDVYKTPSLQAGHDLNKM--	59
QY	57	QOYSNNKDLIEVNEILAEAFNGKNDPLFKPYHYHDTGLLMSACSOEGLDRLGVRVRP--	114
Db	60	GIRLRNGPOLQ-----LSLESLDMQNDLQFKPFHQVGMIDCSSSEKIEIENLRKYOTLL	115
QY	115	----GSDPNLVELTRPEQFRKLAPEGVLQGDPPGKGYEARSAGAGWAHARNALVAAAREA	170
Db	116	DAGIGLEKTNVWLESEDEILAKAP-NFTREQVKGWKGLEFCTDG-GWLAATAKAINALIGFL	173
QY	171	QRMGVKFTVTGTPOGRVVTILIF--ENNNDKGVAVTGDGKIWRARTFICAGASAGQFLDFKN	228
Db	174	QDKGVKFGFG-GAGTFQQPLFAADGKTCIGLETTDGTKYFADKVVLAAGAWSPTLVDEL	232
QY	229	QLRPTAWTLVHIALKPEERALKYNIPIVFNIERGFFPEDEBERGEIKICDEHPGYTNM--	286
Db	233	QCVSKAWVFAHIQLTPKEADAYKNVVDYGEYGFPEFN-EYGVIKVCDEFPFGSRFKL	291
QY	287	---VOSADGTWMSIPEKTOIPKE-----AETRVALLKETMPQLADRFPSFARICWAD	338
Db	292	HQPYGAASPKNISVPSRSHAKHPTDTYDASEVTIRKAIARFLPEFKDKELFNRTMCCTD	351
QY	339	TANREFLIDRHPQYHSLVLCGASGRGFKYLPISIGNLIYDAMEGKVPQKIHELIIKNPDI	398
Db	352	TADANLLICEHPKWNFILATGDSGHSFKLFPNIQKHVVLELLEGSLSQEMAGAWRVRPGG	411
QY	399	AANRWDRDTLGRFGGPNRVMDFHVKW	426
Db	412	DALRS-----RRGAPAK--DLAEMPGW	431

Search completed: May 27, 2005, 13:43:54
Job time : 188.855 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:28:55 ; Search time 2.63539 Seconds

(without alignments)
2331.700 Million cell updates/sec

Title: US-10-622-893A-1

Perfect score: 63

Sequence: 1 MCGSGDDDLAL 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	48	76.2	64	2	033246
2	48	76.2	64	2	Q7T212
3	47	74.6	230	2	Q6H4X1
4	46	73.0	105	2	Q6ZE21
5	46	73.0	392	2	Q6FXT3
6	46	73.0	399	2	Q9NE98
7	46	73.0	951	2	Q6FQ85
8	46	73.0	1039	2	Q9N974
9	45	71.4	586	2	Q6PLR1
10	44	69.8	182	2	Q8L432
11	44	69.8	240	2	Q23696
12	44	69.8	359	2	Q6Z134
13	44	69.8	621	2	Q84X11
14	44	69.8	652	2	Q7XR00
15	44	69.8	718	2	Q91TR1
16	44	69.8	758	2	Q8SAX7
17	44	69.8	997	2	Q6K7F3
18	43	68.3	216	2	Q6EPB1
19	43	68.3	300	2	Q84PB3
20	43	68.3	316	2	Q69X10
21	43	68.3	333	2	Q9S277
22	43	68.3	344	2	Q8RYL3
23	43	68.3	361	2	Q8Z726
24	43	68.3	512	2	Q673G5
25	43	68.3	652	2	Q6YU05
26	43	68.3	708	2	Q8S555
27	43	68.3	894	2	Q9U297
28	43	68.3	925	2	Q86J08
29	42	66.7	64	2	Q30517
30	42	66.7	70	2	Q8GVR6
31	42	66.7	90	2	Q9BY99

32	42	66.7	94	2	Q6K896
33	42	66.7	118	2	Q6UU77
34	42	66.7	135	2	Q69WT7
35	42	66.7	142	2	Q7PWL9
36	42	66.7	203	2	Q7XND1
37	42	66.7	214	2	Q8RYU4
38	42	66.7	215	2	Q6TX11
39	42	66.7	313	2	Q18704
40	42	66.7	316	2	Q6FTB7
41	42	66.7	341	2	Q7FLJ7
42	42	66.7	353	2	Q96XY4
43	42	66.7	417	2	Q8GRJ1
44	42	66.7	455	2	Q6BGW1
45	42	66.7	484	2	Q84Z25

ALIGNMENTS

RESULT 1

O33246	O33246	PRELIMINARY;	PRT;	64	AA.
AC	O33246; Q7D711;				
DT	01-JAN-1998 (TrEMBLrel. 05, Created)				
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)				
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)				
DE	Hypothetical protein.				
GN	OrderedLocusNames=MT2171, RV2111c;				
OS	Mycobacterium tuberculosis.				
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.				
OX	NCBI_TaxID=1773;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=H37RV;				
RX	MEDLINE=9825987; PubMed=9634230; DOI=10.1038/31159;				
RA	Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C.M.,				
RA	Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,				
RA	Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T.,				
RA	Connor R., Davies R.M., Delvin K., Feltwell T., Gentles S., Hamlin N.,				
RA	Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,				
RA	Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,				
RA	Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,				
RA	Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;				
RT	"Deciphering the biology of Mycobacterium tuberculosis from the				
RT	complete genome sequence."				
RL	Nature 393:537-544 (1998).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CDC 1551 / Oshkosh;				
RX	MEDLINE=22206494; PubMed=12218036;				
RX	DOI=10.1126/JB.184.19.5479-5490.2002;				
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,				
RA	Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,				
RA	Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,				
RA	Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,				
RA	Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,				
RA	Fraser C.M.;				
RT	"Whole-genome comparison of Mycobacterium tuberculosis clinical and				
RT	laboratory strains."				
RL	J. Bacteriol. 184:5479-5490 (2002).				
DR	EMBL; BX842578; CAB10702.1; -				
DR	EMBL; AE000516; AAK46454.1; -				
DR	PIR; B70512; B70512.				
DR	TIGR; MT2171; -				
DR	TubercuList; RV2111c; -				
DR	InterPro; IPR008515; DUF797.				
DR	Pfam; PF05639; DUF797; 1.				
SQ	SEQUENCE 64 AA; 6944 MW; 7D958B770D8766E3 CRC64;				

Query Match 76.2%; Score 48; DB 2; Length 64;

Best Local Similarity 80.0%; Pred. No. 1.1;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;


```

QY      2 GGGGDDDDLLA 11
Db      10 GGGGDDDDLLA 19

RESULT 2
Q7TZ12
ID Q7TZ12 PRELIMINARY; PRT; 64 AA.
AC Q7TZ12;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Mb2135C.
GN OrderedLocusNames=Mb2135C;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1763;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248341; CAD96988.1; -.
DR InterPro; IPR008515; DUF797.
DR Pfam; PF05639; DUF797; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 64 AA; 6944 MW; 7D958B770D8766E3 CRC64;

Query Match 76.2%; Score 48; DB 2; Length 64;
Best Local Similarity 80.0%; Pred. No. 1.1;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 GGGGDDDDLLA 11
Db      10 GGGGDDDDLLA 19

RESULT 3
Q6H4X1
ID Q6H4X1 PRELIMINARY; PRT; 230 AA.
AC Q6H4X1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OSJNBa0085K21.17 (Hypothetical protein
DE OSJNBa0023117.33).
GN Names=OSJNBa0085K21.17; Synonyms=OSJNBa0023117.33;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248341; CAD96988.1; -.
DR InterPro; IPR008515; DUF797.
DR Pfam; PF05639; DUF797; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 64 AA; 6944 MW; 7D958B770D8766E3 CRC64;

Query Match 76.2%; Score 48; DB 2; Length 64;
Best Local Similarity 80.0%; Pred. No. 1.1;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 GGGGDDDDLLA 11
Db      10 GGGGDDDDLLA 19

RESULT 4
Q6ZE21
ID Q6ZE21 PRELIMINARY; PRT; 105 AA.
AC Q6ZE21;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein P0495H05.26.
GN Name=P0495H05.26;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248341; CAD96988.1; -.
DR InterPro; IPR008515; DUF797.
DR Pfam; PF05639; DUF797; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 105 AA; 11527 MW; DE2D43D9C75D50CF CRC64;

Query Match 73.0%; Score 46; DB 2; Length 105;
Best Local Similarity 72.7%; Pred. No. 4.3;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 MGGSGDDDDLLA 11
Db      69 LGGGGDDDDSA 79

RESULT 5
Q6FXT3
ID Q6FXT3 PRELIMINARY; PRT; 392 AA.
AC Q6FXT3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to sp|P34223|Saccharomyces cerevisiae YBL058w SHP1.
GN ORFNames=CAGL0A04345g;
OS Candida glabrata CBS138.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=284593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrans P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neugeglise C., Talia E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Niclaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekalia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";

```



```

RL Nature 430:35-44(2004).
DR EMBL; CR380947; CAG57849.1; -.
DR InterPro; IPR006556; DUF_PAF1.
DR InterPro; IPR009060; UBA_like.
DR InterPro; IPR010112; UBX.
DR Pfam; PF00789; UBX; 1.
DR SMART; SM00553; SEP; 1.
DR SMART; SM00166; UBX; 1.
DR PROSITE; PS00033; UBX; 1.
SQ SEQUENCE 392 AA; 43786 MW; 80FF82FCB5E398B CRC64;

Query Match 73.0%; Score 46; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GSGGDDDD 9
Db 106 GSGGDDDD 113

RESULT 6
Q9NE98 PRELIMINARY; PRT; 399 AA.
AC Q9NE98;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein L4803.06.
GN Name=L4803.06;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Masuy D., Purnelle B., Goffeau A., Ivens A.C., Quail M.,
RA Rajadream M.A., Barrell B.G.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL161416; CAB77684.1; -.
KW Hypothetical protein.
SQ SEQUENCE 399 AA; 43458 MW; BB86021C2BBB3E18 CRC64;

Query Match 73.0%; Score 46; DB 2; Length 399;
Best Local Similarity 88.9%; Pred. No. 20;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GSGGDDDD 10
Db 255 GSGGDDDEL 263

RESULT 7
Q6FQ85 PRELIMINARY; PRT; 951 AA.
AC Q6FQ85;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Candida glabrata strain CBS138 chromosome I complete sequence.
GN ORFNames=CAGL0108305g;
OS Candida glabrata CBS138.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=284593;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG Genotevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Karrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikoleki M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekai F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CR380955; CAG60546.1; -.
DR GO; GO:0008415; F:acyltransferase activity; IEA.
DR InterPro; IPR000542; Carn_acyl_trans.
DR Pfam; PF00755; Carn_acyltransf; 1.
SQ SEQUENCE 951 AA; 106462 MW; FB5334D57409D4E1 CRC64;

Query Match 73.0%; Score 46; DB 2; Length 951;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MGGSGDDDDLAL 12
Db 899 IGDEGDDDDLAL 910

RESULT 8
Q9N974 PRELIMINARY; PRT; 1039 AA.
AC Q9N974;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein P1295.14.
GN Name=P1295.14;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Masuy D., Purnelle B., Goffeau A., Ivens A.C., Quail M.,
RA Rajadream M.A., Barrell B.G.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL359773; CAB95222.1; -.
DR HSSP; Q92830; If68.
DR InterPro; IPR001487; Bromodomain.
DR InterPro; IPR011124; Znf_CW.
DR Pfam; PF07496; zf-CW; 1.
KW Hypothetical protein.
SQ SEQUENCE 1039 AA; 109546 MW; 3A91F5A70147E6B3 CRC64;

Query Match 73.0%; Score 46; DB 2; Length 1039;
Best Local Similarity 72.7%; Pred. No. 63;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GSGGDDDDLAL 12

```

```

Db      841 GGGGDDDDVPL 851
|| |||||:|
RESULT 9
Q6PLR1
ID Q6PLR1 PRELIMINARY; PRT; 586 AA.
AC Q6PLR1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Polyphenol oxidase (Fragment).
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA Jukanti A.K., Bruckner P.L., Fischer A.M.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY596266; AF05523.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR008922; Di-copper centre.
DR InterPro; IPR000896; Hemocyanin.
DR InterPro; IPR006311; Tat.
DR InterPro; IPR002227; Tyrosinase.
DR Pfam; PF00264; Tyrosinase; 1.
DR PRINTS; PR00092; TYROSINASE.
DR TIGRFAMs; TIGR01409; TAT signal seq. 1.
DR PROSITE; PS00210; HEMOCYANIN 2; UNKNOWN_1.
DR PROSITE; PS00498; TYROSINASE_2; 1.
FT NON_TER 1
SQ SEQUENCE 586 AA; 63631 MW; D4A1E80D26EE33FE CRC64;

Query Match 71.4%; Score 45; DB 2; Length 586;
Best Local Similarity 80.0%; Pred. No. 47;
Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 3 GGGDDDDLAL 12
|:|||||
Db 36 GTGGDDDLVL 45

RESULT 10
Q8L432
ID Q8L432 PRELIMINARY; PRT; 182 AA.
AC Q8L432;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSJNBb0086108.9 (Hypothetical protein
OSJNBb0086108.4).
GN ORFNames=OSJNBb0086108.4, OSJNBb0086108.9;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,
RA Overton II L.L., Bera J.J., Tsitrin T., Krol M.I., Jarrahl B.B.,
RA Jin S.S., Koo H., Zismann V., Heiao J., Blunt S., Vanaken S.S.,
RA Utterback T.T., Feldblym T.V., Yang Q.Q., Haas B.J., Suh B.B.,
RA Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3]

SEQUENCE FROM N.A.
The Rice Chromosome 10 Sequencing Consortium;
"In-depth view of structure, activity, and evolution of rice
chromosome 10.";
Science 300:1566-1569 (2003).
[4]
RP SEQUENCE FROM N.A.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC105746; RAM92808.1; -.
DR EMBL; AC105746; RAM92811.1; -.
DR EMBL; AF017066; AAP52498.1; -.
DR EMBL; AF017066; AAP52503.1; -.
DR Gramene; Q8L432; -.
KW Hypothetical protein.
SQ SEQUENCE 182 AA; 19521 MW; BDD9842B26EB3F9D CRC64;

Query Match 69.8%; Score 44; DB 2; Length 182;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGSGDDDDLAL 12
|:|||||
Db 1 MCGAGDDDRVEL 12

RESULT 11
O23686 PRELIMINARY; PRT; 240 AA.
ID O23686
AC O23686;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein At2g30480.
GN Name=At2g30480;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Town C.D., Haas B.J., Wu D., Maiti R., Hannick L.I., Chan A.P.,
RA Tallon L.J., Rooney T., Utterback T.R., Vanaken S.E., Feldblym T.V.,
RA White O., Fraser C.M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U93215; AAB63096.1; -.
DR PIR; H84708; H84708.
KW Hypothetical protein.
SQ SEQUENCE 240 AA; 27400 MW; A6AB9BAF4E1CFFCC CRC64;

Query Match 69.8%; Score 44; DB 2; Length 240;
Best Local Similarity 77.8%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGSGDDDD 9
|:|||||
Db 163 VGGTGDDDD 171

RESULT 12
Q6Z134 PRELIMINARY; PRT; 359 AA.
ID Q6Z134
AC Q6Z134;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)

```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein P0455F03.29.
GN Name=P0455F03.29;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
  Clone:P0455F03.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005454; BAC84337.1; -.
DR InterPro; IPR004949; DUF266.
DR Pfam; PF03267; DUF266; 1.
KW Hypothetical protein.
SQ SEQUENCE 359 AA; 39100 MW; C9FDD923301906CB CRC64;

Query Match 69.8%; Score 44; DB 2; Length 359;
Best Local Similarity 66.7%; Pred. No. 40;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MGGSGDDDDLLAL 12
Db 44 VGGGGEDGDLAL 55

RESULT 13
Q84X11 PRELIMINARY; PRT; 621 AA.
AC Q84X11
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Ac2930480/T6B20.25.
GN Name=At2930480/T6B20.25;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eurosidia II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Xiao Y.L., Malik M., Whitelaw C.A., Town C.D.;
RT "Cloning and sequencing of cDNAs for hypothetical genes from
  chromosome 2 of Arabidopsis.";
RL Plant Physiol. 130:2118-2128(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Xiao Y., Smith S.R., Ishmael N., Kumar N., Redman J., Riedmuller S.,
RA Uterback T., Whitelaw C.A., Fraser C.M., Town C.D.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY219097; AAO37184.1; -.
KW Hypothetical protein.
SQ SEQUENCE 621 AA; 71079 MW; A38A7DE170AFC9D3 CRC64;

Query Match 69.8%; Score 44; DB 2; Length 621;
Best Local Similarity 77.8%; Pred. No. 75;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGGSGDDDD 9
Db 490 VGGGGDDDD 498

RESULT 14
Q7XRMO PRELIMINARY; PRT; 652 AA.
AC Q7XRMO

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE OSJNBa0058G03.4 protein.
GN Name=OSJNBa0058G03.4;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Peng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Wang Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL731606; CAB02429.1; -.
DR Gramene; Q7XRMO; -.
SQ SEQUENCE 652 AA; 72720 MW; DF79F807D9962BD4 CRC64;

Query Match 69.8%; Score 44; DB 2; Length 652;
Best Local Similarity 80.0%; Pred. No. 80;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGS GGDDDDLLA 11
Db 21 GGS GGDDDDRA 30

RESULT 15
Q91TRI PRELIMINARY; PRT; 718 AA.
AC Q91TRI
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE T32.
OS Tupaiid herpesvirus 1 (strain 1) (TuHV-1) (Herpesvirus tupaia (strain
  1)).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2.
RX MEDLINE=21211637; PubMed=11312357;
RX DOI=10.1128/JVI.75.10.4854-4870.2001;
RA Bahr U., Darai G.;
RT "Analysis and characterization of the complete genome of tupaia (tree
  shrew) herpesvirus.";
RL J. Virol. 75:4854-4870(2001).
DR EMBL; AF281817; AAK57076.1; -.
DR Pfam; PF06070; Herpes_UL32; 1.
DR PROSITE; PS00583; PFKB_KINASES_1; UNKNOWN 1.
SQ SEQUENCE 718 AA; 74589 MW; 92ED97DAC14EA58F CRC64;

Query Match 69.8%; Score 44; DB 2; Length 718;
Best Local Similarity 88.9%; Pred. No. 89;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGS GGDDDDLL 10
Db 374 GGLGGDDDDL 382

```

Search completed: May 27, 2005, 13:47:24
Job time : 6.63539 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:47:36 ; Search time 126.721 Seconds
(without alignments)
1189.648 Million cell updates/sec

Title: US-10-622-893A-3

Perfect score: 2329

Sequence: 1 AVTKSSLLIVGAGTGTST.....MDFHDKWNTVQYRDISKL 437

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1462099 seqs, 344972447 residues

Total number of hits satisfying chosen parameters: 1462099

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*

6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*

10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*

11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*

12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*

13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*

14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*

15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*

16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*

17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*

18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*

19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2329	100.0	437	17	US-10-622-893A-3
2	2329	100.0	472	17	US-10-622-893A-5
3	1151.5	49.4	440	17	US-10-470-678-9
4	1150.5	49.4	440	17	US-10-470-678-10
5	1149.5	49.4	440	17	US-10-470-678-8
6	625	26.8	437	14	US-10-232-655-3
7	600	25.8	437	14	US-10-232-655-1
8	210.5	9.0	139	16	US-10-767-701-53616
9	184	7.9	384	14	US-10-156-761-14478
10	165	7.1	387	17	US-10-829-427-1
11	165	7.1	385	15	US-10-282-122A-68225
12	160	6.9	433	14	US-10-287-274-412
13	160	6.9	433	15	US-10-282-122A-42701

14	159.5	6.8	381	14	US-10-156-761-9328	Sequence 9328, Ap
15	140	6.0	374	15	US-10-282-122A-55515	Sequence 55515, A
16	136.5	5.9	372	15	US-10-282-122A-43213	Sequence 43213, A
17	135.5	5.8	378	10	US-09-975-719-367	Sequence 367, App
18	135	5.8	420	15	US-10-425-114-54552	Sequence 54552, A
19	127	5.5	372	15	US-10-282-122A-75876	Sequence 75876, A
20	126.5	5.4	433	9	US-09-815-242-13745	Sequence 13745, A
21	126	5.4	370	15	US-10-282-122A-72847	Sequence 72847, A
22	124	5.3	429	15	US-10-425-114-71622	Sequence 71622, A
23	123	5.3	414	16	US-10-437-963-177242	Sequence 177242, A
24	122	5.2	406	17	US-10-485-710-80	Sequence 80, Appl
25	119.5	5.1	391	15	US-10-282-122A-69762	Sequence 69762, A
26	119.5	5.1	818	14	US-10-156-761-14479	Sequence 14479, A
27	113.5	4.9	395	9	US-09-873-880-30	Sequence 30, Appl
28	113.5	4.9	395	15	US-10-424-599-268694	Sequence 268694, A
29	113.5	4.9	852	15	US-10-282-122A-50546	Sequence 50546, A
30	112.5	4.8	317	15	US-10-282-122A-61710	Sequence 61710, A
31	112.5	4.8	405	14	US-10-156-761-9148	Sequence 9148, Ap
32	112.5	4.8	542	16	US-10-437-963-117091	Sequence 117091, A
33	112.5	4.8	626	15	US-10-447-013-27	Sequence 27, Appl
34	112.5	4.8	629	9	US-09-741-669-473	Sequence 473, App
35	112.5	4.8	629	9	US-09-815-242-10405	Sequence 10405, A
36	112.5	4.8	629	15	US-10-282-122A-56758	Sequence 56758, A
37	112	4.8	563	15	US-10-415-302-8	Sequence 8, Appli
38	112	4.8	563	15	US-10-415-302-18	Sequence 18, Appl
39	110.5	4.7	412	9	US-10-416-537-2	Sequence 2, Appli
40	110.5	4.7	632	15	US-10-282-122A-68951	Sequence 68951, A
41	110.5	4.7	797	15	US-10-369-493-9773	Sequence 9773, Ap
42	110.5	4.7	621	15	US-10-282-122A-55707	Sequence 55707, A
43	109.5	4.7	621	15	US-10-282-122A-168785	Sequence 168785, A
44	109	4.7	481	16	US-10-437-963-168785	Sequence 168785, A
45	108.5	4.7	837	15	US-10-282-122A-49792	Sequence 49792, A

ALIGNMENTS

RESULT 1

US-10-622-893A-3

Sequence 3, Application US/10622893A

Publication No. US20050014935A1

GENERAL INFORMATION:

APPLICANT: Yuan, Chong-Sheng

APPLICANT: Datta, Abhijit

APPLICANT: Wang, Yuping

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETERMINATION OF GLYCATED PROTEINS

FILE REFERENCE: 46692001300

CURRENT APPLICATION NUMBER: US/10/622.893A

CURRENT FILING DATE: 2003-07-17

NUMBER OF SEQ ID NOS: 23

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 437

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: 40%-100% identity to the amadoriase

US-10-622-893A-3

Query Match	100.0%	Score	2329	DB	17	Length	437
Best Local Similarity	100.0%	Pred. No.	3.7e-226				
Matches	437	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	AVTKSSLLIVGAGTGTSTALHARGYNTVLDYPVPFSAISAGNDVKNVLSGQYS	60				
Db	1	AVTKSSLLIVGAGTGTSTALHARGYNTVLDYPVPFSAISAGNDVKNVLSGQYS	60				
Qy	61	NNKDEIVNEILLAEAFNGWKNDPLPKFYVHDTGLLMSACSQEGLDLRLGVRRPGEDEPNL	120				
Db	61	NNKDEIVNEILLAEAFNGWKNDPLPKFYVHDTGLLMSACSQEGLDLRLGVRRPGEDEPNL	120				
Qy	121	VELTRPEQFKLAPEGVLOQDFFGWKGYFARSGAGWAHARNALVAAAREAQRMGVKFTVTG	180				

```
Db 121 VELTRPEQFRKLAPEGVLQDPPGWKGYFARSAGAGWAHARNALVAAAREAQRMGVKFTVG 180
QY 181 TPQGRVVTLLIFENNVDKGAVTGDKIWRABERTFLCAGASAGQFLDPFNQLRPTAWTLVHI 240
Db 181 TPQGRVVTLLIFENNVDKGAVTGDKIWRABERTFLCAGASAGQFLDPFNQLRPTAWTLVHI 240
QY 241 ALKPEERALLYKNIPVIFNIERGGFFFPDEERGEIKICDEHPGYTNMVQSADGTMSIPPE 300
Db 241 ALKPEERALLYKNIPVIFNIERGGFFFPDEERGEIKICDEHPGYTNMVQSADGTMSIPPE 300
QY 301 KTOIPKEAETRVALLKETMPOLADRPFSFARICWCADTANREFLIDRHPQYHSLVLCG 360
Db 301 KTOIPKEAETRVALLKETMPOLADRPFSFARICWCADTANREFLIDRHPQYHSLVLCG 360
QY 361 ASGRGFKYLPISGNLIVDAMEGKVPQKIHILIKWNPDIANRNWRDTLGRFGGPNRVMDF 420
Db 361 ASGRGFKYLPISGNLIVDAMEGKVPQKIHILIKWNPDIANRNWRDTLGRFGGPNRVMDF 420
QY 421 HDVKWNTNVQYRDISKL 437
Db 421 HDVKWNTNVQYRDISKL 437

RESULT 2
US-10-622-893A-5
; Sequence 5, Application US/10622893A
; Publication No. US20050014935A1
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong-Sheng
; APPLICANT: Datta, Abhijit
; APPLICANT: Wang, Yuping
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; DETERMINATION OF GLYCATED PROTEINS
; FILE REFERENCE: 466992001300
; CURRENT APPLICATION NUMBER: US/10/622,893A
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric protein
US-10-622-893A-5

Query Match 100.0%; Score 2329; DB 17; Length 472;
Best Local Similarity 100.0%; Pred. No. 4.2e-226;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVTKSSLLIVGAGTGTSTALHLARRGYTNVTVDLPYPVPSAISAGNDVNKVISGGQYS 60
Db 13 AVTKSSLLIVGAGTGTSTALHLARRGYTNVTVDLPYPVPSAISAGNDVNKVISGGQYS 72
QY 61 NNKDETEVNEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQSGLDRLGVRVPGEDPNL 120
Db 73 NNKDETEVNEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQSGLDRLGVRVPGEDPNL 132
QY 121 VELTRPEQFRKLAPEGVLQDPPGWKGYFARSAGAGWAHARNALVAAAREAQRMGVKFTVG 180
Db 133 VELTRPEQFRKLAPEGVLQDPPGWKGYFARSAGAGWAHARNALVAAAREAQRMGVKFTVG 192
QY 181 TPQGRVVTLLIFENNVDKGAVTGDKIWRABERTFLCAGASAGQFLDPFNQLRPTAWTLVHI 240
Db 193 TPQGRVVTLLIFENNVDKGAVTGDKIWRABERTFLCAGASAGQFLDPFNQLRPTAWTLVHI 252
QY 241 ALKPEERALLYKNIPVIFNIERGGFFFPDEERGEIKICDEHPGYTNMVQSADGTMSIPPE 300
Db 253 ALKPEERALLYKNIPVIFNIERGGFFFPDEERGEIKICDEHPGYTNMVQSADGTMSIPPE 312
QY 301 KTOIPKEAETRVALLKETMPOLADRPFSFARICWCADTANREFLIDRHPQYHSLVLCG 360
```

```
Db 313 KTOIPKEAETRVALLKETMPOLADRPFSFARICWCADTANREFLIDRHPQYHSLVLCG 372
QY 361 ASGRGFKYLPISGNLIVDAMEGKVPQKIHILIKWNPDIANRNWRDTLGRFGGPNRVMDF 420
Db 373 ASGRGFKYLPISGNLIVDAMEGKVPQKIHILIKWNPDIANRNWRDTLGRFGGPNRVMDF 432
QY 421 HDVKWNTNVQYRDISKL 437
Db 433 HDVKWNTNVQYRDISKL 449

RESULT 3
US-10-470-678-9
; Sequence 9, Application US/10470678
; Publication No. US20050101771A1
; GENERAL INFORMATION:
; APPLICANT: KOUZUMA, Takuji et al.
; TITLE OF INVENTION: COMPOSITION FOR ASSAYING GLYCATED PROTEINS
; FILE REFERENCE: 1516-0121P
; CURRENT APPLICATION NUMBER: US/10/470,678
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: JP 2001/ 22953
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: JP 2001/ 39796
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: JP 2001/ 24002
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic protein sequence of SEQ ID NO: 3
; NAME/KEY: MUTAGEN
; LOCATION: (372)..(372)
US-10-470-678-9

Query Match 49.4%; Score 1151.5; DB 17; Length 440;
Best Local Similarity 51.0%; Pred. No. 4.2e-107;
Matches 222; Conservative 75; Mismatches 131; Indels 7; Gaps 5;

QY 2 VTKSSLLIVGAGTGTSTALHLARRGYTNVTVDLPYPVPSAISAGNDVNKVIS--SGQY 59
Db 4 LTKSQSLLIVGGTGTSTALHLARRGYTNVTVDLPYPVPSAISAGNDVNKVIS--SGQY 63
QY 60 SNNKDEIEVNEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQSGLDRLGVRVPGED-P 118
Db 64 SKGDDSDSIWKALSVAQAQGLHDPVFPQFCHNTGVSAGSTPKSIKQL-VEDEIGDDID 122
QY 119 NLVELTRPEQFRKLAPEGVLQDPPGWKGYFARSAGAGWAHARNALVAAAREAQRMGVK 178
Db 123 QYTPLNTAEDFRKTMPEGILTGNFPGWKGYFPTGSGWHAHAKAMAAFEESRLGVKFI 182
QY 179 TGTPOGRVVTLLIFENNVDKGAVTGDKIWRABERTFLCAGASAGQFLDPFNQLRPTAWTLV 238
Db 183 TGSPEKVESLIFEDGDVGRGAKTADGKEHRAORTILSAGASAEFFLDFENQIPTAWTLG 242
QY 239 HIALKPEERALLYKNIPVIFNIERGGFFFPDEERGEIKICDEHPGYTNMVQSADGTMSI 297
Db 243 HIQMTPEETKLYKNLPPLFNINQGFMEPEDLHQLKMCDEHPGYCNWVEKPGSKYPQSI 302
QY 298 PEKTOIPKEAETRVALLKETMPOLADRPFSFARICWCADTANREFLIDRHPQYHSLV 357
Db 303 PFAKHQVTEAERMRKQFLKIMPQLADPLVHARICWCADTODRMFLITYHPRHPSLVI 362
QY 358 GCASGRGFKYLPISGNLIVDAMEGKVPQKIHILIKWNPDIANRNWRDTLGRFGGPNRV 417
Db 363 ASGDCGTGYWHITSIGKFSIDCMGTELEERFAKFWWRPEKFTFPGKOPDLDFGADDKI 422
QY 418 MDF--HDVKWNTNVQ 430
```

RESULT 5
US-10-470-678-8
: Sequence 8. Application US/10470678


```
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68225
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-68225

Query Match          7.1%; Score 165; DB 15; Length 385;
Best Local Similarity 23.5%; Pred. No. 1.8e-07;
Matches 101; Conservative 53; Mismatches 161; Indels 114; Gaps 22;

QY      4 KSSLLIVGAGTGTSTALHLARRGYNTVLDPPVPSAISAGNDVNKVISS-----56
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 KSYTAVVGVGVGSATLWELAAQ---QQDVL-----GLEAGAFINLQSSYGGSRIF 55

QY      57 -GOYSNNKDEIEVNEILAEAFNGWK--NDPLFKPYHTDTGLMSACSQEGLDRLGRVR 113
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
56 RQAYWEGSDYLS---LLAEDLGMRELQATSHRPLLYSG-----GLFTG 97

QY      114 PGEDPNLVELTRPEQFRKLAPEGVLQDPPGKGYFARGAGWAHARNALVAAAREA---170
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
98 P-----IRSGVSVGSAAAS-----AKAG-GIAHOR--LTAAEVBARFS 131

QY      171 -----QRMGVKFVTGT-----PQGRVVVTLLIFENNVDKGV---TG 202
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 VFRADENMEAVFQGAFTIAADDARLQMLNQAHAHQMRFGSHVQDITRVESEFLLRLS 191

QY      203 DGKIMRAERTFLCAGAS-AGQFL-DFKNQLRPTAWTLVHIALKPERALYKNIIPV-IFNI 259
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
192 DGQSVLAQKVVLTAGAGLGLIPDLISGLLRPSRVPITYWCAFKSGAGQLFANPAPFLYEL 251

QY      260 ERG--FFFEPEDEERGE--IKICDEHPCYTNMVOSA-DGTMMSIPPEKTOIPKEAETRVRA 314
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
252 EDGRLLYGTPOINDAEPGKI-----GFHNHQQSALDPTQLEPASDAQIEE-----ISA 301

QY      315 LLKETWPLQADRPFSEFARICWCADTANREFLIDRHPQYHSLVLCGASGRGFKYLPISGN 374
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
302 CVSRVFPDLIARYA-SRKCVTYMTTPEAFIIGESKELPSVFVYSCSGHGPKFAFALGS 360

QY      375 LIVDAMEGK 383
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 CLARALAGQ 369
```

RESULT 12

```
US-10-287-274-412
; Sequence 412, Application US/10287274
; Publication No. US20030181408A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THEREOF
; FILE REFERENCE: ELITRA.008DVI
; CURRENT APPLICATION NUMBER: US/10/287,274
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: US 09/711164
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 412
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-287-274-412

Query Match          6.9%; Score 160; DB 14; Length 433;
Best Local Similarity 23.3%; Pred. No. 7.1e-07;
Matches 106; Conservative 49; Mismatches 146; Indels 154; Gaps 24;
```

```
QY      8 LLIIVGAGTGTSTALHLARRGYNTVLDPPVPSAISAGNDVNKVISSGOYNNKDEIE 67
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 IIIIGAGIAGTAGALRCARAG-LSVLLLERABIPG-----SKNLSGRL-----59

QY      68 VNEILAEAFNGWKNDPLFKPYHTDTGLMSACSQEGLDRL---GVRVTPGEDPNLVELT 124
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 YTHALAE-----LLPQHILTAPLERRITHESLSLLTPDGVTT-----96

QY      125 RPEQFRKLAPEG---VLQDPPGKGYFARGAGWAHARNALVAAAREAQMVKFVTG 180
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
97 ----FSSLPQGGESWSLRAFPDW-----LVA---EAEKEGVECI---130

QY      181 TPQGRVVTLIFENNVDKGVAVTGDGKIWAERTFLCAGASAGQFLDFKNQLRPTAWTLVHI 240
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
131 -FCATVDALYEENRGVCGVICGD-DILRARYVVLAEGA-----NSVLAERHGLV--177

QY      241 ALKPERALYKNIIPVFN-----IERGFEFFPEDEERGEI---KICDEHPG---YTNMVQ 288
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
178 -TRPAGEAMALGIKEVLSLETSABERPHLENNEGAALLPFSGRICDDLPGGAPLYTNQQT 236

QY      289 SADG-----TMSIIPPEKTOIPKEAETRVRLIKET-----MPQ--LADRPFS 329
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 LSLGIVCPLESSLTQSRVPASELLTRFKAHPAVRPLIKNTESLEYGAHLVPEGGLHMPVQ 296

QY      330 FARICW-----CADTANREFLIDRHPQYHSLVLCGASGRGPKY----LPSIGNLI 376
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 YAGNGWLLVGDALRSCVNT-----GISVRGMDMALTCGAAAAQTL 336

QY      377 VDAMEGKVPQKTHELIKWNPDIAANRNWEDTIGRF 411
Db      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 ISACQHRPEQNLPPL--YHHNVERSLLM-DVLQRY 368
```

RESULT 13

```
US-10-282-122A-42701
; Sequence 42701, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
```


Query Match	6.0%;	Score 140;	DB 15;	Length 374;
Best Local Similarity	20.7%;	Pred. No. 5.9e-05;		
Matches	90;	Conservative	63;	Mismatches 147; Indels 134; Gaps 22;
Qy	8	LLIIVAGCTGTGTSTALHLARGYTNVTVLDPVPVPSAISAGNDVNKVISSGGSYNNKKDIE	67	
Db	7	LIILIGSVSAGYYATQAG-LKVLIMIDAHLPFSESGSHHGDFTLI---RHAYGEGERY	62	
Qy	68	VNEILAEAFNGWKNDPLFKP-----YYHTGLMSACSQ-----EGL	105	
Db	63	VPLVLRAQTL--W--DELAQKTEDRIPERTGVINLGPQAQSEFLAGVERSAKAFNLDFEKL	118	
Qy	106	DRLGVRVPCGEDPNLVELTRPEQFKL--APBEGVLOQD--FPGWKGYFARSAGCAWAHARN	161	
Db	119	DATGITAR-----WPEITVDDYIGLFEAHSGVLHCETAIKTWIDILAAK--AGCAQLFN	170	
Qy	162	ALVAAAREAQRMGVKFTGTGPQGRVWTLIFENNDVKGAVTGDGKIWRAERTFLCAGASAG	221	
Db	171	CPVEA-----ITHDADGTVVTI-----DGG-YASRLLLVSAG-----	202	
Qy	222	QFLDFKNQLRPTAW---LVHIAUKPERALYKNIPVIENTERGFFFPDDEERGEIKICD	278	
Db	203	-----TWVTRLPLPQIPVRK-----VFS-----WFQSD---GRYSAQN	234	
Qy	279	EHPGYTNMVOSADGTWMSIPEKQTQI-----PKEAETRVRAALKKETMPOLAD	325	
Db	235	KPFAFTGELPNGD-QFYGFPEKQALKIKGGGQAISTPEERN-----LFGAFPQDGS	287	
Qy	326	RPFSFAR-----ICWCADTANREFLIDRHPQYHSLVLGCAGSGRGFKYLPFSI	372	
Db	288	EAFVFLITLPGVGLLYGAACCTVDNTPDEDFIITLPGHNVLLITLCSLGHGKFASVL	347	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - protein search, using sw model

Run on: May 27, 2005, 13:34:21 ; Search time 33.5437 Seconds
(without alignments)
1253.491 Million cell updates/sec

Title: US-10-622-893A-3
Perfect score: 2329
Sequence: 1 AVTKSSLLIVGAGTGTST.....MDPHDVKEWTVQYRDISKL 437

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: Piri:.*
2: Piri:.*
3: Piri:.*
4: Piri:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	488.5	21.0	412	2 T40295	fructosyl amine -
2	465	20.0	433	2 T37605	probable fructosyl
3	192	8.2	390	2 I39975	sarcosine oxidase
4	181	7.8	389	2 JS0671	sarcosine oxidase
5	177.5	7.6	390	2 JC7256	L-pipecolate oxida
6	168	7.2	389	2 T44248	sarcosine oxidase
7	160	6.9	433	1 B65058	fixC protein homol
8	155	6.7	387	2 JU0461	sarcosine oxidase
9	150	6.4	433	2 E91081	hypothetical prote
10	150	6.4	433	2 F85926	hypothetical prote
11	148	6.4	380	2 AB3160	santhopine deamina
12	146.5	6.3	427	2 G83298	conserved hypothet
13	142.5	6.1	857	2 S16133	dimethylglycine de
14	139.5	6.0	364	2 B83078	probable D-amino a
15	138	5.9	376	2 A83905	hypothetical prote
16	137.5	5.9	372	2 A85668	sarcosine oxidase-
17	137.5	5.9	372	2 E90808	sarcosine oxidase-
18	136.5	5.9	372	2 JC5371	probable sarcosine
19	136.5	5.9	382	2 G75122	sarcosine oxidase,
20	136.5	5.9	416	2 AD3075	sarcosine oxidase
21	136.5	5.9	416	2 E98211	sarcosine oxidase
22	135.5	5.8	372	2 F90159	sarcosine oxidase,
23	134.5	5.8	825	2 F95963	probable dehydroge
24	132	5.7	371	2 F83487	hypothetical prote
25	128.5	5.5	382	2 F71008	probable sarcosine
26	127	5.5	372	2 A10637	opine catabolism h
27	127	5.5	372	2 A70019	probable oxidoredu
28	126.5	5.4	433	2 A10882	thiamin biosynthes
29	126	5.4	652	2 AH2245	

30	125.5	5.4	417	2 G98200	sarcosine oxidase
31	125.5	5.4	417	2 AB3086	sarcosine oxidase
32	125.5	5.4	439	2 E82983	probable oxidoredu
33	125.5	5.4	442	2 D97382	hypothetical prote
34	125.5	5.4	442	2 AC2600	oxidoreductase ord
35	123.5	5.3	437	2 T45533	agaE protein (impo
36	116.5	5.0	452	2 G95356	probable oxidoredu
37	114	4.9	396	2 AF2835	sarcosine oxidase
38	114	4.9	413	2 A97613	sarcosine oxidase
39	113	4.9	420	2 AC3284	D-amino-acid dehyd
40	113	4.9	703	2 T15503	hypothetical prote
41	112.5	4.8	629	1 BVECOA	glucose inhibited
42	112.5	4.8	629	2 C91214	glucose-inhibited
43	112.5	4.8	629	2 D86060	glucose-inhibited
44	112	4.8	405	2 I40887	sarcosine oxidase
45	111.5	4.8	441	2 F98306	agaE protein (impo

ALIGNMENTS

RESULT 1

T40295
fructosyl amine - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T40295
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, March 1998
A;Reference number: Z21918
A;Accession: T40295
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-412 <MOO>
A;Cross-references: UNIPROT:O43029; EMBL:AL022071; PIDN:CAA17815.1; GSPDB:GN00067; SPDB:
A;Experimental source: strain 972h-; cosmid c354
C;Genetics:
A;Gene: SPDB:SPBC354.15
A;Map position: 2

Query Match 21.0%; Score 488.5; DB 2; Length 412;
Best Local Similarity 30.1%; Pred. No. 5e-32;
Matches 130; Conservative 78; Mismatches 183; Indels 41; Gaps 12;

Qy	2	VTKSSLLIVGAGTGTSTALHARRGYTNVTLDVYPVPSAISAGNDVKNKVISSGOYSN	61
Db	1	MVKNTSVIIVGAGVGLSAALELTGRGYTIKILDRAPPVIDGSSVDANRIIRS-DYAD	59
Qy	62	NKDEIEVNEILAEAFNGWKNDPLFKPYHDTGLMSA-----CSQGLDLGLYVR	113
Db	60	-----AVYCSMGIDALEEWNRTNLFKEQFYGSLMFVGRDNVEYRDMLENLTOMGV---	111
Qy	114	PGEDPNLVELTRPEOKLAPEGLQDFPGWGYFARSGAGWAHARNALVAAREAQRM	173
Db	112	-----SAAKPQTTEERLKLFPKWI--GELNDGAGYANFSSGWNABQSVVNYLAHA	164
Qy	174	GKVFVTGTPQGRVVTILFENNVDKGAVTGDKGIWRAERTFLCAGASAGQL-DFKNQLRP	232
Db	165	GVSFISG-PBGTVEELITEENVVKGVRTTGT-AYMAEKLIFATGAWTASLLPNDHTRFLA	222
Qy	233	TATLVHIALKPERALYKNIPIVFNIERGFPEPDEERGEIKCDHPHYTNMVQSDG	292
Db	223	TGQFVAYIKLTPEYIRFLNPNVYLPDFTGTGYIFPPPTPDYGLYFARHGYGFTRMNLKSG	282
Qy	293	TWMSIPEKTQI-----PKAEATRVALLKETM-POLADRPFSFARICWCADTANREFLID	347
Db	283	KVESVPPKPLVSPILPKEABLDLRNLRNLTQTYEEISQRFYKTRICYTDTADAEEVFD	342
Qy	348	RHPQYSHLVLCGASGRGFKYLPISIGNLIVDAMEGKVPQKIHILIKW---NPDIANRNW	404
Db	343	YHPDYENLNVCTGSGHGFFKFFPLGLKYSIGCMFRELEELPKKWKWKENLEPAALDHS	402
Qy	405	RDTLGRFGGPNR	416

Db 403 R-----AGPSR 408 ||:|

RESULT 2

T37605

probable fructosyl amino acid oxidase - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T37605

R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Saunders, D.; Harris, D. submitted to the EMBL Data Library, October 1999

A:Reference number: Z21729

A:Accession: T37605

A:Status: preliminary; translated from GB/EMBL/DDBU

A:Molecule type: DNA

A:Residues: 1-433 <CD>

A:Cross-references: UNIPROT:Q9UTW9; EMBL:AL132667; PIDN:CAB59618.1; GSPDB:GN00066; SPDB:

A:Experimental source: strain 972h-; cosmid c139

C:Genetics:

A:Gene: SPDB:SPAC139.04C

A:Map position: 1

Query Match 20.0%; Score 465; DB 2; Length 433;

Best Local Similarity 30.1%; Pred. No. 4.5e-30;

Matches 130; Conservative 80; Mismatches 174; Indels 48; Gaps 13;

QY 5 SSSLLVGGTGTCTALHLLAR-RGYTNVTLDPYPVPSAISAGNDVNKVISGQYSNNK 63

Db 2 SRTIVVGGCVGLSTAVELAKNHSFDNTIAIDAEVPSPSSMAANDINKIVRP-EYADLK 60

QY 64 DEIEVNEILAAEFNGKNDPLFKPYHDTGLLMSACSQEGLDRLGVRVRPGP----- 116

Db 61 -----YMKLALAEKWRNDPELSSVYFEGRLST-----ISKDPRARFDEVAQRNL 109

QY 117 -----DPNLVELTPEQFRKLAP-----EGVLQDFFPGWKGYFARSGAGWAHARNALVAAR 168

Db 110 KLLGDSALINLSSEIEIRKKYPSLFSNPLRSMDQA-----VNEHAGYANSAASKLEL 165

QY 169 EAQRMGVKFTVTPQGR-----VV-----TLIFENNVDKGAVTGDKIWAERTFLCAGASA 220

Db 166 KARELGVEFVFG-KAGFKFKFVNNHSETDIDKNDNHSVQTEGDTIYHADTILLAVGAYL 224

QY 221 GQFLDFKNQLRPTAWTLVHIALKPERALYKNIPVFNIERGFFPEDEERGIKICDEH 280

Db 225 NAYLNTSHRVCAGLPAHQLTDEEFKYKNPILFPDCAVAFPPYPVKLIKLASTG 284

QY 281 PGYTNMVQS---ADGTWMSIPF---EKTQIPKEAETRVLRALLKETMPQLADRPFSFARIC 334

Db 285 YEYVCNVETDYDENSKVSIPIHSGPSKSLPKYALIQMRFRFLDPLDLADRLINTKMC 344

QY 335 WCADTANREFLIDRHPQYHSLVLCGASGRGFKYPSIGNLIYDVAMEGKVPQKIHELK 394

Db 345 WISDTEADNPLDKVPQFDNVFVANGDSGHAFKPLNIGRYIAQRIILGDLSEWKDAWR 404

QY 395 NPDIAANR-NWR 405

Db 405 REDDKASELKWR 416

RESULT 3

I39975

sarcosine oxidase (EC 1.5.3.1) - Bacillus sp.

C:Species: Bacillus sp.

C>Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004

C:Accession: I39975

R:Suzuki, K.; Sagai, H.; Inamura, S.; Sugiyama, M. J. Ferment. Bioeng. 77, 231-234, 1994

A:Title: Cloning, sequencing, overexpression in Escherihia coil of a sarcosine oxidase-

A:Reference number: I39975

A:Accession: I39975

A:Status: preliminary; translated from GB/EMBL/DDBU

A:Molecule type: DNA

A:Residues: 1-390 <RES>

A:Cross-references: UNIPROT:P40859; GB:D16521; NID:G984787; PIDN:BAA03967.1; PID:G984788

C:Genetics:

A:Gene: sox

C:Superfamily: sarcosine oxidase

C:Keywords: oxidoreductase

Query Match 8.2%; Score 192; DB 2; Length 390;

Best Local Similarity 23.2%; Pred. No. 8.1e-08;

Matches 97; Conservative 68; Mismatches 162; Indels 92; Gaps 22;

QY 8 LLIIVGAGTGTSTALHLLARRGYTNVTLDPYPVPSAISAGNDVNKVISGQYSNNKDEIE 67

Db 7 VIVVGAGSMGAAGYQLAKQGVKTLIV-DAFDPHTNGSHHGDTIRHA-YGEGREYVP 64

QY 68 VNEILAAEFNGKNDPLFKPYH---DTGLLMSACSQEGLDRLGVRVRPGEDPNLVEL 123

Db 65 -----LALRSQELWYE-LEKETHHKIFTKTGLV-----FGPK---GESAFVAET 105

QY 124 TRPEQFRKLAPGVLQGD-----FPGW---KGYFA-RSGAGWAHARNALVAAAREAORM 173

Db 106 MEAAKEHSLTVD-LLEGDEINKRWPGITVPEYNAIFEPNSGVLFSENCIRAYRELAER 164

QY 174 GVKFTVTPQGRVVTLLIFENNVDKGAVTGDKIWAERTFLCAGASAGQFLDFKNQLRPT 233

Db 165 GAKVLTHT---RVEDPDISPDSVK-IETANGS-YTADKLIVSMGAWNSKLLSKLN----- 214

QY 234 ANTLVHIALKPERALYKNIPVFNIERGFFPEDEERGIKICDRHPGYTNMVQSADGT 293

Db 215 -----LDIPLQ-----YRQVV-----GPFESDESKYSNDI--DFFPGF--MVEVPNGI 253

QY 294 MMSIPF-----EKTQIPKEAETRVLRALLKETMPQLADRPFSF 330

Db 254 YGFPSPGCGGLKLGVTGQKIDPDTINREFGVPEDESNLRAFLLEYMPG-ANGELKR 312

QY 331 ARICWCADTANREFLIDRHPQYHSLVLCGASGRGFKYLPSPIGNLIYD-AMEGKVPQKI 388

Db 313 GAVCMYKTLDEHFDIHLHPEHNSVIAAGFSGHGFSSGVGEVLSQLALTGKTSHDI 371

RESULT 4

JS0671

sarcosine oxidase (EC 1.5.3.1) precursor - Streptomyces sp.

C:Species: Streptomyces sp.

C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Jun-2000

C:Accession: JS0671; PS0345

R:Suzuki, K.; Ogishima, M.; Sugiyama, M.; Inouye, Y.; Nakamura, S.; Inamura, S. Biosci. Biotechnol. Biochem. 56, 432-436, 1992

A:Title: Molecular cloning and expression of a Streptomyces sarcosine oxidase gene in Str

A:Reference number: JS0671; MUID:92330009; PMID:1368326

A:Accession: JS0671

A:Molecule type: DNA

A:Residues: 1-389 <SUZ>

A:Cross-references: GB:D10623; NID:G217023; PIDN:BAA01473.1; PID:G217024

A:Experimental source: strain KB210-8S

A:Accession: PS0345

A:Molecule type: protein

A:Residues: 2-31;122-143;230-259;268-283 <SUZ1>

A>Note: the authors translated the initiation codon GTG for residue 1 as Val

C:Comment: This enzyme catalyzes the oxidation of sarcosine to form formaldehyde, glycine

C:Superfamily: sarcosine oxidase

C:Keywords: oxidoreductase

F:2-389/Product: sarcosine oxidase #status experimental <SAR>

F:11-16/Region: nucleotide binding #status predicted

Query Match 7.8%; Score 181; DB 2; Length 389;

Best Local Similarity 22.9%; Pred. No. 6.4e-07;

Matches 100; Conservative 59; Mismatches 150; Indels 128; Gaps 24;

QY 8 LLIIVGAGTGTSTALHLLARRGYTNVTLDPYPVPSAISAGNDVNKVISGQYSNNKDEIE 67

Db 7 VIVIGGGMGSAAAHLSARG-ARVLGLEKF-----GPNHRGSSHG 47

RESULT 7
B65058
fixC protein homolog b2766 - Escherichia coli (strain K-12)
N;Contains: probate quinone reductase (EC 1.6.5.-)
C;Species: Escherichia coli
C;date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002
C;Accession: B65058
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.;
A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997	
A:Title: The complete genome sequence of Escherichia coli K-12.	
A:Reference number: A67420; MUID:97426617; PMID:9278503	
A:Accession: B65058	
A>Status: preliminary; nucleic acid sequence not shown; translation not shown	
A:Molecule type: DNA	
A:Residues: 1-433 <BLAT>	
A:CROSS-references: GB:AE000360; GB:U00096; NID:g2367157; PIDN:AACT5808.1; PID:g1789125;	
A:Experimental source: strain K-12, substrain MG1655	
C:Superfamily: fixC protein	
C:Keywords: oxidoreductase	
Query Match	6.9%; Score 160; DB 1; Length 433;
Best Local Similarity	23.3%; Pred. No. 3.9e-05;
Matches	106; Conservative 49; Mismatches 146; Indels 154; Gaps 24;
QY 8 LLIVGAGTGTSTALHLARGYTNVTLDPYPVPSAISAGNDVNKVISSQYSNKKDEIE 67	
DB 17 IIIIGAGIAGTACALRCARAG-LSVLLLERAEIPG-----SKNLSGRL----- 59	
QY 68 VNEILAEAFNGWKNDPLFKPYVYHDTGLMSACSQGLDRL--GVRVRPGEDPNLVELT 124	
DB 60 YTHALAE-----LLPQHLTAPLERRITHESLSLLTPDGVT----- 96	
QY 125 RPEQFKLAPEG-----VLQDFFGWKGYFARSGAGWAHARNALVAAAREAQRMGVKFTVG 180	
DB 97 -----FSSLPQGGESWSVLARFDPW-----LVA---EAEKEGVBCI-- 130	
QY 181 TPQGRVVTLLIFENNVDKGVATGDKIWAERTFLCAGASAGQFLDPFNQLRPTATWLVHI 240	
DB 131 -PGATVDALYEENGRVCGVICGD-DILRARYVYVLAEGA-----NSVLAERHGLV-- 177	
QY 241 ALKPEERALKNIPVIFN-----IERGFFPEDEERGEI---KICDEHPG-----YTNMVQ 288	
DB 178 -TRPAGEAMALGKIVLSLETSIAIERFHLNENGAALLFSGGICDDLPFGAFLYTNQOT 236	
QY 289 SADG-----TMMSSIPPEKTIQPKAEATRVALLKET-----MPQ--LADRPFS 329	
DB 237 LSLGIVCPSSLTSQSRVASELUTRFKAHPAVRPLIKNTESLEYGAHLVPEGGLSMPVQ 296	
QY 330 FARICW-----CADTANREFLDRHPQYHSLVLCGASGRGFKY-----LPSIGNLI 376	
DB 297 YAGNGWLLVGDALRSCVNT-----GISVRGMDMALTGAAQAQTL 336	
QY 377 VDAMEGKVPQKIHELKWNPDIAANRNWRDTLGRF 411	
DB 337 ISACQHPQNUPFL--YHNHVERSLW-DVLQRY 368	
RESULT 8	
JU0461	
sarcosine oxidase (EC 1.5.3.1) - Bacillus sp.	
C:Species: Bacillus sp.	
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004	
C:Accession: JU0461	
R:Koyama, Y.; Yamamoto-Otake, H.; Suzuki, M.; Nakano, E.	
Agric. Biol. Chem. 55, 1259-1263, 1991	
A:Title: Cloning and expression of the sarcosine oxidase gene from Bacillus sp. NS-129	
A:Reference number: JU0461; MUID:91291342; PMID:1368683	
A:Accession: JU0461	
A:Molecule type: DNA	
A:Residues: 1-387 <KOY>	
A:CROSS-references: UNIPROT:P23342; GB:D10553; GB:D01067; NID:g216326; PIDN:BAA01410.1;	
A:Experimental source: strain NS-129	
C:Comment: This enzyme catalyzes the oxidation of sarcosine to generate formaldehyde, g	
C:Superfamily: sarcosine oxidase	
C:Keywords: oxidoreductase	
Query Match	6.7%; Score 155; DB 2; Length 387;
Best Local Similarity	22.8%; Pred. No. 8.5e-05;
Matches	94; Conservative 62; Mismatches 178; Indels 78; Gaps 22;
QY 8 LLIVGAGTGTSTALHLARGYTNVTLDPYPVPSAISAGNDVNKVISSQYSNKKDEIE 67	

Db 7 VIVGAGSGMGAAGYILAKQGVKTLIV--DSFPDPHTNGSHGDTRIIRHA-YEGREYVP 64	
QY 68 VNEILAEAFNGWKNDPLFKPYVYHDTGLIM-----SACSQGLDRLGVVRVPGEDPNLV 121	
DB 65 F-ALQAQELWYELEKETHHK-IFTQTGVLVGPKGSAFVSETM-----EAAANH 112	
QY 122 ELTRPEQFKLAPEGVLQDFFGWK-----YFA-RSGAGWAHARNALVAAAREAQRM 173	
DB 113 SLEH-ELF-----EGQLTD--RWAGVEVPDNYEAFIPNSGVLFSENCIQYRELAEAH 164	
QY 174 GVKFVTGTP--QGRVVTLIFENNVDKGVATGDKIWAERTFLCAGASAGQFLDPFNQLR 231	
DB 165 GATVITYTPVEDFEVTEDLVTIKTAGSYT-----ANKLVVSGAMNSKLSKLD---- 214	
QY 232 PTATWLVHIALKPE-----ERALKYK-----IPV-IPNIERGFFFS-PDEERGEIKIC 277	
DB 215 -----VEIPLOPYQVGVFFECDEAKYSNNAHYAFWVEVENGLIYYGFPFSGGLKI- 267	
QY 278 DEHPGYTNMVQADGTMMSIPPEKTIQPKAEATRVALLKETMPQLADRPFSFARICWCA 337	
DB 268 ----GVHSYGQIIDPDTINREFGAY---PEDEANLRKFLEQYMPG-ANGELKKGAVCMYT 319	
QY 338 DTANREFLDRHPQYHSLVLCGASGRGFKYLPISIGNLIVD-AMSGKVPQKI 388	
DB 320 KTPDEHFEVIDLHPKYSNVAIAAGFSGHGFKFSVVVGETLAQLATTGKTEHDI 371	
RESULT 9	
E91081	
hypothetical protein Ecs3621 [imported] - Escherichia coli (strain O157:H7, substrain RIN	
C:Species: Escherichia coli	
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004	
C:Accession: E91081	
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;	
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.	
DNA Res. 8, 11-22, 2001	
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom	
A:Reference number: A99629; MUID:21156231; PMID:11258796	
A:Accession: E91081	
A>Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-433 <HAY>	
A:CROSS-references: UNIPROT:Q8X7T8; GB:BA000007; PIDN:BA37044.1; PID:g13363092; GSPDB:G	
A:Experimental source: strain O157:H7, substrain RIMD 0509952	
C:Genetics:	
A:Gene: Ecs3621	
C:Superfamily: fixC protein	
Query Match	6.4%; Score 150; DB 2; Length 433;
Best Local Similarity	22.9%; Pred. No. 0.00026;
Matches	104; Conservative 49; Mismatches 148; Indels 154; Gaps 24;
QY 8 LLIVGAGTGTSTALHLARGYTNVTLDPYPVPSAISAGNDVNKVISSQYSNKKDEIE 67	
DB 17 IIIIGAGIAGTACALRCARAG-LSVLLLERAEIPG-----SKNLSGRL----- 59	
QY 68 VNEILAEAFNGWKNDPLFKPYVYHDTGLMSACSQGLDRLGVVRVPGEDPNLVELTRPE 127	
DB 60 YTHALAE-----LLPQHLTAPLERRITHESL-----SLTTPD 92	
QY 128 ---QFKLAPEG-----VLQDFFGWKGYFARSGAGWAHARNALVAAAREAQRMVKFTVG 180	
DB 93 CATTFSSLPQGGESWSVLARFDPW-----LVA---EAEKEGVBCI-- 130	
QY 181 TPQGRVVTLLIFENNVDKGVATGDKIWAERTFLCAGASAGQFLDPFNQLRPTATWLVHI 240	
DB 131 -PGATVDALYEENGRVCGVICGD-DILRARYVYVLAEGA-----NSVLAERHGLV-- 177	
QY 241 ALKPEERALKNIPVIFN-----IERGFFPEDEERGEI---KICDEHPG-----YTNMVQ 288	
DB 178 -TRPAGEAMALGKIVLSLETSIAIERFHLNENGAALLFSGGICDDLPFGAFLYTNQOT 236	


```
QY 289 SADG-----TWMSIPFEKTIQPKAEATVRALLKET-----MPQ--LADRPFS 329
Db 237 LSLGIVCLPSLTOSRVPASELLARFKTHPAVRPLIKNTLSLEYGAHLVPEGGHLSMPVQ 296
QY 330 FARICW-----CADTANREFLIDRHPQYHSLVLCGASGRGPKY-----LPSIGNLI 376
Db 297 YAGNGWLLVGDALRSCVNT-----GISVRGMDTALTGAQAAQTLL 336
QY 377 VDAMEGKVPQKIHELKWNPDIAANRNWRDTLGRF 411
Db 337 ISACQHREPQNLFFL--YHNNVERSLW--DVLQRY 368

RESULT 10
F85926
hypothetical protein ygcN [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: F85926
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F85926
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-433 <STO>
A;Cross-references: UNIPROT:Q8X7T8; GB:AE005174; NID:g12517227; PIDN:AAG57874.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ygcN
C;Superfamily: fixC protein

Query Match 6.4%; Score 150; DB 2; Length 433;
Best Local Similarity 22.9%; Pred. No. 0.00026;
Matches 104; Conservative 49; Mismatches 148; Indels 154; Gaps 24;

QY 8 LLIVGAGTGTCTALHLARRGYTNVTLDPPVPSPAISAGNDVNVKVISSQGYNNKDEIE 67
Db 17 IIIIGAGIAGTACALRCARAG-LSVLLLEAEIPG-----SKNLGGRL----- 59
QY 68 VNEILAEANGWKNDFLPKPYHYDHTGLMSACSQEGLDLGRVVRPGEDPNLVELTRPE 127
Db 60 YTHALAE-----LLPQFHLTAPLERRITHESL-----SLITPD 92
QY 128 ---QFRKLAPEG---VLOGDFPGWKGYFARSAGWAHARNALVAAAREAQRMGVKFTVG 180
Db 93 CATTFSSLPQGGESWSVLRARFPW-----LVA---EAEKGVVECI-- 130
QY 181 TPQGRVVTLIFENNVDKGAVTGDKIWAERTFLCAGASAGQFLDFKNQLRPTAWTLVHI 240
Db 131 -PGATVDALYEENRGVGVICGD-DILRARYVVLAEGA-----NSVLAERHGLV-- 177
QY 241 ALKPEERALYKNTPVIFN-----IERGFFPEDEERGEI---KICDEHPG-----YNNMVQ 288
Db 178 -TRPAGEAMALGIKEVLSELSAIEERFHLNENGAALFSGGICDDLPGCAFLYTNQOT 236
QY 289 SADG-----TWMSIPFEKTIQPKAEATVRALLKET-----MPQ--LADRPFS 329
Db 237 LSLGIVCLPSLTOSRVPASELLARFKTHPAVRPLIKNTLSLEYGAHLVPEGGHLSMPVQ 296
QY 330 FARICW-----CADTANREFLIDRHPQYHSLVLCGASGRGPKY-----LPSIGNLI 376
Db 297 YAGNGWLLVGDALRSCVNT-----GISVRGMDTALTGAQAAQTLL 336
QY 377 VDAMEGKVPQKIHELKWNPDIAANRNWRDTLGRF 411
Db 337 ISACQHREPQNLFFL--YHNNVERSLW--DVLQRY 368

RESULT 11
AB3160
santhopine deaminating protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 16-Aug-2004
C;Accession: AB3160
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AB3160
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-380 <KUR>
A;Cross-references: UNIPROT:Q8UKU5; GB:AE008687; PIDN:AAL45696.1; PID:g17743424; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: agaE
C;Superfamily: Sarcosine oxidase

Query Match 6.4%; Score 148; DB 2; Length 380;
Best Local Similarity 22.2%; Pred. No. 0.00031;
Matches 90; Conservative 51; Mismatches 151; Indels 114; Gaps 18;

QY 6 SLLIVGAGTGTCTALHLARRGYTNVTLDPPVPSPAISAGNDVNVKVISSQGYNNKDE 65
Db 12 SSAVIIGAGTGTGVLARRG-IQVTILNDGP-FANGASGRSLWNSARSRSEPHQ 69
QY 66 IEVNEI-----LAEE-----AFNG-----WKND-----PLFKPYHYDHTGLMSACSQEGLD 106
Db 70 LRMAGIDRYRTLAAENPDVEWLRFDGGLTWDSDERNEIDAAVREHVSLEYDA-----Q 123
QY 107 RLG-----VVRPGEDPNLVELTRPEQFKLAPGCVLQDPPGKGYFARSAGWAHARNA 162
Db 124 RLSAGDVARVTPGIDAGI-----TPQGAI-----FNPGEQGVLDLPTL 161
QY 163 LVAAREAQRMGVKFTVGTGQGRVVTLIFENNVDKGAVTGDKIWAERTFLCAGASAGQ 222
Db 162 IRVLLLEEFALGGVLTDOGAARV---MLEGGRVAGTAETAGVYLRADAVVLATGPAVK 218
QY 223 FLDFKNQLRPTAWTLVHIALKPEERALYKNIPVIFNIE-----GFFPEDEERGEIK 275
Db 219 MVGESQIIIGDG--TPIALLVQTKPLAHLRAVLNTPRVAVRPAFGSGSLDADWA--- 272
QY 276 ICDEHPGYTNMVOSADGTMSPFETQTPKEAETRVALLKTMFQLADRPFSFARICW 335
Db 273 -ADE-----GVTVRADGT-----EIDDTIVAEILLVEAAK----- 301
QY 336 CADTANREFLIDRHPQYHSLVLCGASGRGKYLPSIGNLIVDAME 381
Db 302 -----VMEGNPOLEVASIGVGG-----KPIPGDGEPTVGAIK 333

RESULT 12
G83298
conserved hypothetical protein PA2776 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: G83298
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brj
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: AB2950; MUID:20437337; PMID:10984043
A;Accession: G83298
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-427 <STO>
A;Cross-references: UNIPROT:Q9I067; GB:AE004705; GB:AE004091; NID:g9948851; PIDN:AAG0616
```

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA2776

C;Superfamily: hypothetical protein HI0499

Query Match 6.3%; Score 146.5; DB 2; Length 427;

Best Local Similarity 23.7%; Pred. No. 0.00049;

Matches 98; Conservative 59; Mismatches 164; Indels 93; Gaps 23;

QY 6 SSSLIVGAGTGTSTALHARRGYTNVTLDPYPVPSAISAGNDVNKVISSGQYSNNKDE 65

DB 30 TDVCVIGAGTGLSTALFLLENGF-KVTVLEAAKVGFGAS-GRNGGQIVNS--YERDIDV 85

QY 66 IEVN-----ETLAEEAFNGKNDPLFKPKYHTDTGL-----LMSACSQEGLDRLGVR 111

DB 86 IERTVKREAQLGEMAFEGGR---IIRERVARYGIGQCDLKDGGVFAAFTEKQMDHLRAQ 142

QY 112 VRPGE--DNLVELTRPEOFKRLAPGVLOGDFPGWKGYFARSGAGWAHARNALVAARE 169

DB 143 KQLWERYGHNQLEIMDAKRIR-----EVVATDNYIG--GMLMSG-GHIHPLNLALGEAAA 195

QY 170 AORMGVKFVTGTPOGRVVTLIFENNNDVKGAVTGDGKIWAERTFLCAGASAGQFLDFKNQ 229

DB 196 VESLGRRIYEQSPATRI-----ERGASPVVHTPOGRV---KAKPIVVAGNA--YLN----- 241

QY 230 LRPTAWTLVHIALKPEERALKYKNIP-----VI-----FNIERGFFPEDEERGEIKICDEHP 281

DB 242 -----GLVPELAA--KSMPCGTQVITTEPLNEELAHSLLPQD-----YCVEDC 282

QY 282 GYTNMVQSADGTWMSI-----PFKTKQIPKEAETRVALLKETWMPQLADPPFPFARICWCA 337

DB 283 NYLLDYRLSGDKRLYIGGGVYIGARDPADIEAIRPKMLKTPPQKDKVIDFA--W-- 337

QY 338 DTANREFLDRHPQ-----VHSLVLGCGASGRGFKYLPSTGNLIVDAMEGK 383

DB 338 -TGNFLLTSLRPLQVGRIGDNIYISQ-----GCSGHGVVYTHVAGRVLAELRQ 386

RESULT 13

S16133

dimethylglycine dehydrogenase (EC 1.5.99.2) - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C;Accession: S16133

R;Lang, H.; Polster, M.; Brandsch, R.

Eur. J. Biochem. 198, 793-799, 1991

A;Title: Rat liver dimethylglycine dehydrogenase. Flavinylolation of the enzyme in hepatoc

A;Reference number: S16133; MUID:91266966; PMID:1710965

A;Molecule type: mRNA

A;Residues: 1-857 <LAN>

A;Cross-references: UNIPROT:Q63342; EMBL:X55995; NID:g56688; PIDN:CAA39468.1; PID:g56688

C;Keywords: flavoprotein; oxidoreductase

Query Match 6.1%; Score 142.5; DB 2; Length 857;

Best Local Similarity 20.3%; Pred. No. 0.0028;

Matches 98; Conservative 74; Mismatches 160; Indels 151; Gaps 23;

QY 9 LIVGAGTGTSTALHARRGYTNVTLDPYPVPSAISAGN-----DVNKV 53

DB 46 VIIGGCGVGSYLAHLKAGMRDVILLE-----KSELTAGSTWHAAGLTTFYFHPGINLKKI 101

QY 54 ISSGQYSNNKDETVENELAEAFN--GWKNDPLFKPYHDTGLLMSACSQEGLDRLGVR 111

DB 102 -----HYDSIKLVERLEEETGOVVG-----FHQPSIRLATTPERVDFEKYQ 143

QY 112 V-RPGEDPNLVELTRPEOFKRLAP-----EGVLOGDFPGWKGYFARSGAGWAHARNALVA 166

DB 144 MTRTNWHTAQYIIEPKIHELFPULLNMDKILAGY-----NPGDGHIDPYSLTMAL 195

QY 167 AREAORMGV--KF---VTG-----TQGRVVTLIFENNNDVKGAVTGDGKIWA 209

DB 196 ATGARKYGVLLKYPAPYTSLSKPRPDGTWDETTPQGSV-----RA 234

QY 210 ERTFLCAG---ASAGQFLDFKNQLRPTAWTLVHIALKPEERALKYKNIPVIFNIE----- 260

DB 235 NRIVNAAGFWAREVGVKMGICLDHPLIPVQHQYVVTSTIPEVKALKRELPVLRLDLEGSYYLR 294

QY 261 ---RCGFFPEDEERGEIKICDE-----HPGYTNMVQSADGTWMSIPEKTKQIPKEAETR 311

DB 295 QERDGLLFGPYESQKMKLQASWVAHGVPPGFGKELFESDLDDRI-----TEHVEAAMEM 348

QY 312 VRALLKETWMPOLADRPFSFARICWCADTANREFLDRHPQYHSLVLGCGASGRGFKYLP 371

DB 349 VVVLKADLIINIVNGPIIYS-----PDILP---MVGPHQGVNRYVAIG-FGYGIHAGG 399

QY 372 IGNLIVD-AMEGKVPQKTHELIKWNPDIAANRNWRDITLGRFGGPNRVMDFDHVKWNTNVQ 430

DB 400 VGKYLSDWLHGEFP---FDLIELDPN-----RYG-----KWTITQ 432

QY 431 YRD 433

DB 433 YTE 435

RESULT 14

B83078

probable D-amino acid oxidase PA4548 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 16-Aug-2004

C;Accession: B83078

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Bro

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: B83078

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-364 <STO>

A;Cross-references: UNIPROT:P33642; GB:AE004868; GB:AE004091; NID:g9950785; PIDN:AAG07936

A;Experimental source: strain PA01

C;Genetics:

C;Gene: PA4548

C;Superfamily: Sarcosine oxidase

Query Match 6.0%; Score 139.5; DB 2; Length 364;

Best Local Similarity 24.6%; Pred. No. 0.0014;

Matches 103; Conservative 44; Mismatches 165; Indels 107; Gaps 19;

QY 5 SSSLIVGAGTGTSTALHARRGYTNVTLDPYPVPSAISAGNDVNKVISSGQYSNNKD 64

DB 2 SERVVVVGAGVIGLLLTARELALAGLRVTLVERGESGREASWAGGIVSPLYPMWYSP--- 58

QY 65 EIEVNEILAEAFNGKNDPLFKPYHDTGLLMSACSQEGLDRLGVVRPFGEDPNL--- 120

DB 59 -----AVTALAHWSQD--FYP-----ALGQRLLEDTELGL-----DPEVHTVG 92

QY 121 ---VEL-----TRPEQFRKLAPEGVLOGDFPGWKGYFARS--GAGWAHARNA 162

DB 93 LYWLDDDDQTEALQWARKHTRP---LKEVPIEEAAYAAPGLGAGFORAVVMGSAVANRNP 149

QY 163 LVAAREAQRMGVKFTVTPQGRVVTLIFENNNDVKAVTGDGKIWAERTFLCAGASAGQ 222

DB 150 RLARSRLASLQOPANLELHEQTEVRGWLDRGVRVGVATSRGEI-RGDKVLLAAGAWSG 208

QY 223 FLDPKNQLRPTAWTLVHIALKPEERALKYK---NIPVIFNIERGFFPEDEERGEIKICD 278

DB 209 L-----LKLGLGLELPVVPVK--GOMLLYKCAADFLPRMV-LAKGRVAIPRD--GHILIG 259

QY 279 --EHPGYTNMVQSADGTWMSIPEKTKQIPKEAETRVALLKETWMPQLAD-----R 326

DB 260 TLEHSG-----FDKTP--TDEAQESLRASAAELLPDLADMQPVVAHWAGLR 302

QY 327 PFSFARICWCADTANREFLDRHPQYHSLVLGCGASGRGFKYLPSTGNLIVDAMEGKVP 385

Search completed: May 27, 2005, 13:48:43
Job time : 35.5437 sec8

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:38:36 ; Search time 51.2473 Seconds

(without alignments)

975.171 Million cell updates/sec

Title: US-10-622-893A-3

Perfect score: 2329

Sequence: 1 AVTKSSLLIVGAGTGTST.....MDPHDVKEWTVQYRDISKL 437

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 538906 seqs, 114359116 residues

Total number of hits satisfying chosen parameters: 538906

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pap.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pap.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pap.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pap.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pap.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pap.*
- 7: /cgn2_6/ptodata/2/paa/US11_NEW_COMB.pap.*
- 8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	215	9.2	459	US-60-643-717-13558	Sequence 13558, A
2	181.5	7.8	440	US-11-097-143-4818	Sequence 4818, Ap
3	177.5	7.6	390	US-10-990-328A-13084	Sequence 13084, A
4	177.5	7.6	445	US-10-990-328A-13083	Sequence 13083, A
5	169	7.3	387	US-10-990-477-6	Sequence 6, Appli
6	167	7.2	389	US-10-534-583-1	Sequence 1, Appli
7	165	7.1	387	US-10-990-477-2	Sequence 2, Appli
8	148.5	6.4	774	US-10-990-328A-13788	Sequence 13788, A
9	148.5	6.4	866	US-10-990-328A-13787	Sequence 13787, A
10	142	6.1	388	US-10-526-324-372	Sequence 372, App
11	130.5	5.6	386	US-10-526-324-29	Sequence 29, Appli
12	124.5	5.3	530	US-60-669-241-28797	Sequence 28797, A
13	115.5	5.0	389	PCT-IB03-06509-1503	Sequence 1503, Ap
14	115	4.9	505	US-11-031-175-9986	Sequence 9986, Ap
15	108	4.6	796	US-10-450-763-48529	Sequence 48529, A
16	108	4.6	879	US-10-204-639-58	Sequence 58, Appli
17	106.5	4.6	822	US-60-643-717-17883	Sequence 17883, A
18	105.5	4.5	822	US-60-643-717-7212	Sequence 7212, Ap
19	104.5	4.5	339	US-11-122-943-10	Sequence 10, Appli
20	104.5	4.5	339	US-11-122-943-12	Sequence 12, Appli
21	104	4.5	515	US-11-097-143-42765	Sequence 42765, A
22	100.5	4.3	102	US-60-655-875-133298	Sequence 133298, A
23	99.5	4.3	340	US-11-122-943-18	Sequence 18, Appli
24	98.5	4.2	849	US-60-643-717-18183	Sequence 18183, A
25	98.5	4.2	850	US-60-643-717-17113	Sequence 17113, A

ALIGNMENTS

RESULT 1

US-60-643-717-13558
; Sequence 13558, Application US/60643717

; GENERAL INFORMATION:

; APPLICANT: Abad, Mark S.

; TITLE OF INVENTION: Genes and Uses for Plant Improvement

; FILE REFERENCE: 38-21(53629)A

; CURRENT APPLICATION NUMBER: US/60/643,717

; CURRENT FILING DATE: 2005-01-12

; NUMBER OF SEQ ID NOS: 19247

; SEQ ID NO 13558

; LENGTH: 459

; TYPE: PRT

; ORGANISM: ASPERGILLUS NIDULANS FGSC A4

US-60-643-717-13558

Query Match 9.2%; Score 215; DB 8; Length 459;

Best Local Similarity 19.5%; Pred. No. 1.5e-10;

Matches 82; Conservative 34; Mismatches 84; Indels 220; Gaps 10;

Qy 3 TKSSLLIVGAGTGTSTALHLAGRYNTVTLD-PYPVPSAISAGDNVKNKVISSGQYSN 61

Db 5 SKSIPITAIIGDAGFLSTALHLVQNGYTDITVLEQDEKIPPPYSAANYLNKIV----- 57

Qy 62 NKDEIEVNEILABEAFNGWKNDFLFXPHYHDTGLLMSACSOGLDLRLGVVRPGEENLV 121

Db 58 -REYY-----EDP--- 64

Qy 122 ELTRPEQFKLAPEGVLQDFPGWKGYFARSAGAGAHARNALVAAAREAQRMVKFVTGT 181

Db 65 ----- 64

Qy 182 PQGRVVTLIPENNNDVKAGVTGDKIWAERTFLCAGASAGQFLDFKNQLRPTAWLVHIA 241

Db 65 ----- 64

Qy 242 LKPEERALYKNIPIVFIENERGFFFPDEERGEIKICDEHPGYTMVMSADGTWMSIPFEK 301

Db 65 -----FYNLTIV-----LCFMSGGYIN-TDKATGVSHS-PFPV 95

Qy 302 TQ-----TPKEATRVALLKETMQLADRPSPARICWCADTANREFLIDRHP-QYHS 354

Db 96 NOEASGGLFAEDETIRKLLQOTLPALANRPLVLKSLCWFADTKDSDFIIDFVPGSKGS 155

Qy 355 LVLGCGASGRGFKYLPISIGNLIVDAMEGKVPQKIHILIKWNPDIANRNWDTLGRGCGP 414

Db 156 VVIESADSGHGKPFVIGSVTSLLSKRHTPSNIH---SKPNLRANMGQSTRGNYGLP 212


```

; PRIOR APPLICATION NUMBER: JP 2003-387975
; PRIOR FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: JP 2004-184960
; PRIOR FILING DATE: 2004-06-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 387
; TYPE: PR1
; ORGANISM: Bacillus sp. strain KS-11A
US-10-990-477-6

Query Match          7.3%   Score 169; DB 6; Length 387;
Best Local Similarity 21.4%; Pred. No. 1.5e-06;
Matches      89; Conservative    69; Mismatches 172; Indels     86; Gaps    20;

Qy      8 LLIVGACTGTCTSTALHLARGYTNVTVLDPYPVPSAISAGNDVNKVISSGOYSNNKDEIE 67
Db      7 VVVVGSGMGAAGYILAKQGKVTKLLV-DAPDPPHTGGSHHGDTRIIRHA-YGEGREYVP 64

Qy      68 VNEILASEAFNGWKNDPLFKPYHDTGLL-----MSACSQEGL-----D 106
Db      65 P-ALRAQLWYELENETHNK-IFTKTGVLVFGPKGBSDFVAETWEAAAEHSLTVDLLEGD 122

Qy      107 RLGRVRVRPGEDPNLVELTRPEQFRLAPEGVTLQDDFFCPGWKGYPARSAGWAHARNALVAA 166
Db      123 EINTR-WFG-----ITVENYAI-----PEPNSGVLFSENCIRS Y 157

Qy      167 AREAQRMGVKFVTGTTPGRRVVTVLIFENNVDKAVGTGDGKIWRABERTFLCAGASAGQLDF 226

```

107	Qy	RLGVRVRPGEDPNLVELTRPEQFRKLAPEGVLQGDFFGWKGYFARSAGAGWAHARNALVAA	166
118	Db	-----VDLLEGSEINKRWP-GVTPE--NYNAIFEK-N-SGVLTFSENCIRAY	159
167	Qy	AREAQRMGVKFVTCTPQGRVVTLIFENNVDK-----GAVTGDGKIWAERTFLCAGASAG	221
160	Db	-----RELAEANGAKVLITYP---VEDPEIADFVKIQAYGSF-----ASKLIVSMGAWNS	209
222	Qy	QFLDFKNQLRPTAWTLVHIALKPEERALYKNIPVIFNIERGFPEPEERGE-----	273
210	Db	KLLSKLN-----IEIPLQP-----YRQW-----GPFECDEKKYSNTHGYPAF	247
274	Qy	-IKI-----CDEHPGYTNMVQSADGTMSPPEKTOIPKEAETRVALLKET	319
248	Db	MVSPFGIYYGFSFGCGLKIGVHTYQKIDPDTINREF---GIYPEDENIRKFLETY	304
320	Qy	MPQLADPFFSARICWCAADTANREFLDHRPQHSLVLGCCASGRGPKYLESIGNLIYVD-	378

Db 305 MPG-ATGELKSGAVCYTKTPDEHFVLDLHPQFSNVAIAAGFSGHGKFSVVGETLSQL 363
Qy 379 AMEGKVPQKI 388
Db 364 AVTGKTEHDI 373

RESULT 7
US-10-990-477-2
; Sequence 2, Application US/10990477
; GENERAL INFORMATION:
; APPLICANT: KIKKOMAN CORPORATION
; APPLICANT: FURUKAWA, Keisuke
; APPLICANT: KAJIYAMA, Naoki
; TITLE OF INVENTION: MODIFIED SARCOSINE OXIDASES, MODIFIED SARCOSINE OXIDASE GENES,
; TITLE OF INVENTION: AND METHODS FOR PREPARING THE MODIFIED SARCOSINE OXIDASES
; FILE REFERENCE: 261714US0
; CURRENT APPLICATION NUMBER: US/10/990,477
; CURRENT FILING DATE: 2004-11-18
; PRIOR APPLICATION NUMBER: JP 2003-387975
; PRIOR FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: JP 2004-184960
; PRIOR FILING DATE: 2004-06-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-990-477-2

Query Match 7.1%; Score 165; DB 6; Length 387;
Best Local Similarity 21.4%; Pred. No. 3.3e-06;
Matches 89; Conservative 68; Mismatches 173; Indels 86; Gaps 20;

Qy 8 LLIVGAGTGTSTALHARRGYNTVTLDPYPVPSAISAGNDVKNVYISSGQISNNKDEIE 67
Db 7 VIVVGAGSMGMAAGYLAQKQGVKTLV-DAPDPFHTGSHGGTRIIIRA-YGEGREYVP 64
Qy 68 VNEILAEAFNGWKNDPLFKPYHYDHTGLL-----MSACSOEGL-----D 106
Db 65 F-ALRAQELWYLENEHTNK-IFTKTGLVFGPKGESDFVAETMEAAAHSLLVDLEGD 122
Qy 107 RLGVVRPGEDPNLVELTRPEQKLAPEGLQGDPPGKGYFARSAGAGWAHARNALVAA 166
Db 123 EINTR-WPG-----ITVPENYNAI-----FEPNSGVLFSENCIRSY 157
Qy 167 AREAORMGVKFTGTGTPQGRVVTLLIFENNVDKGVATGDKIWAERTFLCAGASAGQFLDP 226
Db 158 RELAVAKGAKILTYT---RVEDFEVSQDQVK-IQTANGS-YTADKLIVSGAWNSKLSK 212
Qy 227 KQQLRPTAWTLVHIALKP-----EERALKN---IPV-IFNTERGFFPDEERGE 273
Db 213 LN-----LDIPLQPRQVGVGFDSEAKISNDVDYPAFVVEVPKGIYGFPSFGG- 262
Qy 274 IKICDHPGTYNNVQSDGTMMSIPPEKTQIPKEAETRVALLKETMPLADRPFSFARI 333
Db 263 ---CGLKIGVHTYQQIDPDTINREFGAYQ---EDSSNRDRDFLEKYMPE-ANGELKRGAA 315
Qy 334 CWCADTANREFLDHRYPOVHSLVLCGASGRGFKYLPSTGNLIVD-AMEGKVPQKI 388
Db 316 CMYTKTPDEHFVLDTHPEHSNVEVAAGFSGHGKFSVVGEVLSQLATTGKTEHDI 371

RESULT 8
US-10-990-328A-13788
; Sequence 13788, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES

; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13788
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-13788

Query Match 6.4%; Score 148.5; DB 6; Length 774;
Best Local Similarity 20.9%; Pred. No. 0.0025;
Matches 98; Conservative 73; Mismatches 176; Indels 121; Gaps 22;

Qy 9 LIVGAGTGTSTALHARRGYNTVTLDPYPVPSAISAGN-----DVKV 53
Db 53 VLIIGCGVGVSLAYHLAKAGMKDVVLE-----KSELTAGSTWHAAGLTTYFHPGINKKI 108
Qy 54 ISSGQYSNNKDEIEVNEILAEAFN--GKNDPLFKPYHYDHTGLLSACSQEGDLRLGVR 111
Db 109 -----HYDSIKLYEKLKEETGQVVG-----FHQPGSIRLATTPVRVDFEKYQ 150
Qy 112 V-RPGEDPNLVELTRPEQPKLAP---EGVLQGDPPGKGYFARSAGAGWAHARNALVAA 166
Db 151 MTRTGWATEQYLIIEPKIQEMFPLNNMKVLAGLY-----NPGDHIDPYSUTMAL 202
Qy 167 AREAORMGVKFTGTGTPQGRVVTLLIFENNVDKGVATGDKIWAERTFLCAG---ASAGQF 223
Db 203 AAGARKCGALLKVPAP---VTSLKARSDGTWDVETPOGSM-RANRIVNAAAGFAWREVGKM 258
Qy 224 LDFKNQLRPTAWTLVHIALKPEERALKYKIPVIFNIE-----RGFFFPDEERGEI 274
Db 259 IGLEHLPIPVQHYVVTSTIPEVKALKRELPLVRLDEGSYYLRQERDGLLFGYESQEKM 318
Qy 275 KICDE-----HPGYTNMV--QSDGTMMSIPPEKTQIPKEAETRVALLKETMPLADR 326
Db 319 KVQDSWTVNGVPPGKELFESDLRIMEHI-----KAAMWVPLKADIINVNG 370
Qy 327 PFSFARICWADTANREFLDHRYPOVHSLVLCGASGRGFKYLPSTGNLIVD-AMEGKVP 385
Db 371 PITYS-----PDILP---MVGPHQGVRYNVAIG-FGYGIIHAGGKGYLSLWILHGEPP 421
Qy 386 QKIHELKKNPDIAANRWDRDLGRFGGNRVMDPHDKVETWVQYRD 433
Db 422 ---FDLIELDPN-----RYG-----KWTTOYTE 442

RESULT 9
US-10-990-328A-13787
; Sequence 13787, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13787
; LENGTH: 866
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-13787

Query Match 6.4%; Score 148.5; DB 6; Length 866;
Best Local Similarity 20.9%; Pred. No. 0.0003;
Matches 98; Conservative 73; Mismatches 176; Indels 121; Gaps 22;


```
Db          360 DGKTDK 365
|
|
RESULT 11
US-10-526-324-29
; Sequence 29, Application US/10526324
; GENERAL INFORMATION:
; APPLICANT: Imanaka, Takayuki
; APPLICANT: Atomi, Haruyuki
; TITLE OF INVENTION: METHOD OF TARGETED GENE DISRUPTION, GENOME OF
; TITLE OF INVENTION: HYPERTHERMOSTABLE BACTERIUM AND GENOME CHIP USING
; FILE REFERENCE: 490051.401USPC
; CURRENT APPLICATION NUMBER: US/10/526,324
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: PCT/IB2003/003597
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: JP 2002-319011
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 2167
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Thermococcus kodakaraensis KOD1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (414542)..(414542)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (786890)..(786890)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (786907)..(786907)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (786944)..(786946)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (839139)..(839139)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1128488)..(1128488)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1128499)..(1128499)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1128505)..(1128506)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1128517)..(1128518)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1128539)..(1128540)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1561400)..(1561400)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1561477)..(1561477)
```

```
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1561545)..(1561545)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1767941)..(1767941)
; OTHER INFORMATION: n is a or c or g or t.
US-10-526-324-29

Query Match      5.6%; Score 130.5; DB 6; Length 386;
Best Local Similarity 21.0%; Pred. No. 0.0037;
Matches 93; Conservative 65; Mismatches 164; Indels 121; Gaps 23;

QY 4 KSSLLIIVGAGTGWGTSTALHLARRGYTNVVLDPVPFSAISAGNDVNKVISSGGYSNNK 63
Db 8 EKSEITIIGGGIVGVTTIAHELAKRG-EVTVVIE-----KRFIGSGSTFRCGTGIROQFNDE 62
QY 64 DEIEVNEILAEAFNGWKNNDPLFKPY-----YHDTGLLLMSACSQEGLD----- 106
Db 63 ANVQMKRSVE-----LWKYSEYGFPPQQTGYLFLLYDDEEVETFKRNAIQN 112
QY 107 RLGVVRPGEDNLVELTRPEQFRKLAP-EGVLQGDFFGKGYFARSGAGNAHARNALVA 165
Db 113 KFGVPTR-----LITPEAKEIVPLLDISEVVAASWNPDTGKASPFHSTAKFAL-- 161
QY 166 AAREAQRMGKVEVTGTPQGRVVTLIFENNVDKGAVTGDKIWRABRTFLCAGASAGQFLD 225
Db 162 ---HAEEFGAKLVEYT---EVKDFIIEGGEIKGLKTSRGTI-----KTGLVNVATNA---- 207
QY 226 FKNQLRPTAATLVHIALKPEERALKYNIPIVNIERGFEEFPEDEERGEIK--ICDEHPGY 283
Db 208 WAKLINAMAGIRTKPIEP-----YKHQAVI-----TQPIKKGSVKPWVIFRYGH 253
QY 284 TMMVQSADGTM-SIPFEKTQIPKAEATRVALLKETMPQLADRRPFSFARICWCAATANR 342
Db 254 AVLTQTSHGGIIGVGVE-----EGPTYD---LNPVTEFLREVSYFTKII----PALR 300
QY 343 BELI-----DRHP-----QYHSLVLCGAGSGPKYLPISIGNLIVDAM-EGK 383
Db 301 ELLILRTWAGYYAKTPDSNPAIGKIEELSDYYIAAGFSGHGFMMAPAVAEVMDLITKGR 360
QY 384 VPQKIHELKKNPDIAANRNMRD 406
Db 361 T-----DLPAA-----WYD 369

RESULT 12
US-60-669-241-28797
; Sequence 28797, Application US/60669241
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; APPLICANT: Baum, James A
; APPLICANT: Gilbertson, Larry A
; APPLICANT: Kovalic, David K
; APPLICANT: LaRosa, Thomas J
; APPLICANT: Lu, Maolong
; APPLICANT: Munyikwa, Tichifa R. I.
; APPLICANT: Roberts, James K
; APPLICANT: Wu, Wei
; APPLICANT: Zhang, Bei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CONTROL OF INSECT INFESTATION IN PLAN
; FILE REFERENCE: 38-21(53596)
; CURRENT APPLICATION NUMBER: US/60/669,241
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60560842
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 60565632
; PRIOR FILING DATE: 2004-04-27
; PRIOR APPLICATION NUMBER: 60579062
; PRIOR FILING DATE: 2004-06-11
; PRIOR APPLICATION NUMBER: 60603421
```

```

; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 5809
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1503
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Mycobacterium paratuberculosis
PCT-IB03-06509-1503

Query Match          5.0%; Score 115.5; DB 1; Length 389;
Best Local Similarity 22.2%; Pred. No. 0.081;
Matches 53; Conservative 41; Mismatches 96; Indels 49; Gaps 19;

QY      8  LLIVGAGTGTSTALHLARGYNTVTLDPYPVPSAISAGNDVNKVISSQYSNKKDEIE 67
      ::::| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      74  IVVIGAGVCGLAHAHLSRRGHDDVVVLEKGQDPFGQSAG-----LARIPIAHRRRESLC 128

QY      68  VNEILAEAFNGWKNDPLPKFPYYHDTGLLMSACSQSG-----LDRLGVVRPVGDPNVL 123
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      129  RLAFAPARAGQWRWEAFGVRLLGSEGFIAGAAGAADGVAQAMERAGA----- 175

QY      124  TRPEQRFKLAPEGVLQ-----GPPGWMKGVFARSAGAGWAHARNALVAAAREAQRMGVKFT 179
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      176  ----AFSRLDRDGIARIRIPADVFWETGTFDPLG-GSLRIRALTTLAR----RVVIR--- 224

QY      180  GTPQGRVVTLIFENNVDKGAVTGDGKIMWAERTFLCAGASAGQF-----LDFKNQLRPT 233
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      225  ---RGEVSVSA---DNGSTVLADGTVLRADRVLCAGVATPKLFGPLGVDFVPHTFT 276

RESULT 14
US-11-031-175-9986
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/11/031,175
; CURRENT FILING DATE: 2005-01-08
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9986
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-11-031-175-9986

Query Match          4.9%; Score 115; DB 7; Length 505;
Best Local Similarity 21.9%; Pred. No. 0.13;
Matches 92; Conservative 54; Mismatches 148; Indels 126; Gaps 24;

QY      2  VTKSSLLIVGAGTGTSTALHLARGYNTVTLDPYPVPSAISAGNDVNKVISSQYS- 60
      : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      115  VPAAKVVIIIGGINGLALAYNLARAGETDVVVLERGILCAGASGRGGGVRMQWGTPLS 174

QY      61  --NNKDEIEVNEILAE-APNGW-----KNDPLFKPYHDTGLLMSACSQEGLD 107
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      175  VELAKRSIELMKFARELGINVLWLRQGGYIFLAKTAPVAQLERNVSL-----HNR 225

QY      108  LGVVRPVGDPNVLTRPEQRFKLA-----EGVLQGDPPGHWKGYFARS 152
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      226  FGVPTR-----LITPEARGIVPGITMKDCLIASYNPEDGVI---FP-WPFL----- 268

QY      153  GAGWAHARNALVAAAREAQRMGVKFTGTPTQGRVVTLIFENNVDKGAVTGDGKIWAERT 212
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      269  ---WGYAQG-----CQKRGVRVETYT-----DVTGFEISGGQVRKVKT 304

QY      213  -----FLCAGA---SAGQFLDFKNQLRPTAWTLVHIALKPEERALYKNIPIFNIE 260
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 305 RGDIACTVTLAAGAWSPQVAKLADVLPNEPHR-----HEILSTEPKDFLG-PLVSVLD 359
QY 261 RGFFFEPPDEERGEI--KICD-BHPGYTNMVQSGADGTMMSIPFEKTOIPKEAETRVRALEIK 317
Db 360 SGLYFS-QSMRGEIVGGMGDAKEPAGLNM---GSTLRFV-----SREAQALM 402
QY 318 ETWPLQADRPFFARIC--W--CAD-TANREFLIDRHPOYHSLVLGCGASGRGFKYLPISI 372
Db 403 EQLPEV-----GHVKVLRWQAGCYDVTDPNNPILGRTPLGLDNLQMSGFVGHGFMMAVAV 457

RESULT 15
US-10-450-763-48529
; Sequence 48529, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 48529
; LENGTH: 796
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (413)..(443)
; OTHER INFORMATION: 2Fe-2S ferredoxins iron-sulfur binding region proteins.
; OTHER INFORMATION: domain identified by EMATRIX, accession number BL00197A, p-value=
; OTHER INFORMATION: 8.342e-14, raw score of 18.23
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (18)..(113)
; OTHER INFORMATION: ParB-like nuclease domain identified by Pfam, accession name
; OTHER INFORMATION: ParBc, E-value=4.8e-20, Pfam score of 80.0
US-10-450-763-48529

Query Match 4.6%; Score 108; DB 6; Length 796;
Best Local Similarity 20.0%; Pred.No.1;
Matches 87; Conservative 51; Mismatches 118; Indels 178; Gaps 21;
QY 11 VGAGTGTSTALHLARRGYTNVTLPDPYPVPSAISAGDNVNVKVISSGOYSNNKDELEV-- 68
Db 77 IGKG-----SSSLKRLKGLVPTCLE-----GTRNQRIATIRHNRARGHQITA 122
QY 69 -NEILAEAFNGWKNDPLPKPYHYDHTGLMSACSQEGLDRL----- 108
Db 123 MSEIVRELSQLGWDNDKIGKELGMDSEVLRLKQINGLQELFADROYRSRAWTLKLAQRSA 182
QY 109 -----GVRVR-----PCGEDPNLVELTPEQF 129
Db 183 AASNVDSEFLANRDESYGKGSANTDSGIAVRAPASATGKTQHHPGPMVMVMVMDQAF 242
QY 130 RKLAP--EGVLQDFFPGWGYFAR-----SGAG-----WAHARNALVAAAREAQ 171
Db 243 KKLNPHEHPVLHSD-QGWQYRMYRYQNIKEHCGGVTPIMSMRWL-AKNRPQADV--- 297
QY 172 RMGVKPVTTGPQGRV-----VTLIFFENDVKGAVTGDGKIWR-----AERTF 213
Db 298 ---VIYNVTPQDVIFADEWRNYPVTLVAENNVTEGFIA--GRLTRELLAGVPLASRTV 352
QY 214 LCAGASAGQFLDFKNQLRPTAWTLVHLKPEERALKPIPVIFNIERGEF---FPEPDEE 270
Db 353 MTCGPA--PYMDWVEQ-----EVKAL-----GVTRFFKEKFFFT--- 384

QY 271 RGEIKICDEHPGYTNMVQSGADGTMMSIPFEKTOIPKE-----AETRVALLKETWPLQAD 325
Db 385 -----VAEAATGSLKFTKIQPAREFYAPVGTTLLEAESNNVDDVA- 425
QY 326 RPFSPARICWCADT 339
Db 426 ---ACRAGVCGCCKT 437
Search completed: May 27, 2005, 13:51:55
Job time : 53.2473 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:47:36 ; Search time 4.92964 Seconds
(without alignments)
1189.648 Million cell updates/sec

Title: US-10-622-893A-4

Perfect score: 90

Sequence: 1 KGELEGLPIPNPLRTG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1462099 seqs, 344972447 residues

Total number of hits satisfying chosen parameters: 1462099

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA: *

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	17	17 US-10-622-893A-4	Sequence 4, Appli
2	90	100.0	472	17 US-10-622-893A-5	Sequence 5, Appli
3	59	65.6	146	9 US-09-923-304-9	Sequence 9, Appli
4	58.5	65.0	326	16 US-10-865-978-37	Sequence 37, Appl
5	58.5	65.0	338	16 US-10-467-243-26	Sequence 26, Appl
6	58.5	65.0	349	16 US-10-467-243-8	Sequence 8, Appli
7	58.5	65.0	368	14 US-10-270-223-4	Sequence 4, Appli
8	58.5	65.0	473	15 US-10-239-652A-6	Sequence 6, Appli
9	58.5	65.0	473	15 US-10-239-652A-8	Sequence 8, Appli
10	58.5	65.0	473	15 US-10-415-232-4	Sequence 4, Appli
11	58.5	65.0	669	9 US-09-213-888-25	Sequence 25, Appl
12	58.5	65.0	669	9 US-09-328-877A-25	Sequence 25, Appl
13	58.5	65.0	669	16 US-10-653-497-25	Sequence 25, Appl

14	58.5	65.0	669	16	US-10-653-517-25	Sequence 25, Appl
15	58.5	65.0	669	16	US-10-653-496A-25	Sequence 25, Appl
16	58.5	65.0	669	17	US-10-653-676A-25	Sequence 25, Appl
17	58.5	65.0	991	16	US-10-865-978-34	Sequence 34, Appl
18	56.5	62.8	253	16	US-10-624-909-20	Sequence 20, Appl
19	56.5	62.8	361	14	US-10-006-780-8	Sequence 8, Appli
20	56.5	62.8	383	9	US-09-883-096-5	Sequence 5, Appli
21	55.5	61.7	570	17	US-10-800-350-386	Sequence 386, App
22	51.5	57.2	135	16	US-10-624-909-192	Sequence 192, App
23	51.5	57.2	181	16	US-10-624-909-86	Sequence 86, Appl
24	51.5	57.2	196	16	US-10-624-909-160	Sequence 160, App
25	51.5	57.2	200	16	US-10-624-909-48	Sequence 48, Appl
26	51.5	57.2	212	16	US-10-624-909-68	Sequence 68, Appl
27	51.5	57.2	233	16	US-10-624-909-172	Sequence 172, App
28	51.5	57.2	236	16	US-10-624-909-56	Sequence 56, Appl
29	51.5	57.2	237	16	US-10-624-909-66	Sequence 66, Appl
30	51.5	57.2	237	16	US-10-624-909-102	Sequence 102, App
31	51.5	57.2	237	16	US-10-624-909-124	Sequence 124, App
32	51.5	57.2	237	16	US-10-624-909-126	Sequence 126, App
33	51.5	57.2	237	16	US-10-624-909-188	Sequence 188, App
34	51.5	57.2	238	16	US-10-624-909-52	Sequence 52, Appl
35	51.5	57.2	238	16	US-10-624-909-78	Sequence 78, Appl
36	51.5	57.2	238	16	US-10-624-909-84	Sequence 84, Appl
37	51.5	57.2	238	16	US-10-624-909-88	Sequence 88, Appl
38	51.5	57.2	238	16	US-10-624-909-104	Sequence 104, App
39	51.5	57.2	238	16	US-10-624-909-118	Sequence 118, App
40	51.5	57.2	238	16	US-10-624-909-132	Sequence 132, App
41	51.5	57.2	238	16	US-10-624-909-146	Sequence 146, App
42	51.5	57.2	239	16	US-10-624-909-38	Sequence 38, Appl
43	51.5	57.2	239	16	US-10-624-909-44	Sequence 44, Appl
44	51.5	57.2	239	16	US-10-624-909-60	Sequence 60, Appl
45	51.5	57.2	239	16	US-10-624-909-62	Sequence 62, Appl

ALIGNMENTS

RESULT 1

US-10-622-893A-4
; Sequence 4, Application US/10622893A
; Publication No. US20050014935A1
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong-Sheng
; APPLICANT: Datta, Abhijit
; APPLICANT: Wang, Yuping
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: DETERMINATION OF GLYCATED PROTEINS
; FILE REFERENCE: 46692001300
; CURRENT APPLICATION NUMBER: US/10/622, 893A
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 40%-100% identity of the second bacterial leader sequence
US-10-622-893A-4

Query Match 100.0%; Score 90; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 4,7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KGELEGLPIPNPLRTG 17

Db 1 KGELEGLPIPNPLRTG 17

RESULT 2

US-10-622-893A-5
; Sequence 5, Application US/10622893A
; Publication No. US20050014935A1

; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong-Sheng
; APPLICANT: Datta, Abhijit
; APPLICANT: Wang, Yuping
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: DETERMINATION OF GLYCATED PROTEINS
; FILE REFERENCE: 46692001300
; CURRENT APPLICATION NUMBER: US/10/622, 893A
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric protein
US-10-622-893A-5

Query Match 100.0%; Score 90; DB 17; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGEGLPIPNPLRTG 17
|||||
Db 450 KGEGLPIPNPLRTG 466

RESULT 3
US-09-923-304-9
; Sequence 9, Application US/09923304
; Patent No. US20020081612A1
; GENERAL INFORMATION:
; APPLICANT: KATZ, RUTH
; APPLICANT: JIANG, FENG
; TITLE OF INVENTION: DETECTION AND DIAGNOSIS OF SMOKING RELATED CANCERS
; FILE REFERENCE: UTSC:658US
; CURRENT APPLICATION NUMBER: US/09/923,304
; CURRENT FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-923-304-9

Query Match 65.6%; Score 59; DB 9; Length 146;
Best Local Similarity 59.3%; Pred. No. 0.32;
Matches 16; Conservative 0; Mismatches 1; Indels 10; Gaps 2;

QY 1 KGEGLPIPNPLRTG 17
|||||
Db 114 KGEGLPIPNPLRTG 140

RESULT 4
US-10-865-978-37
; Sequence 37, Application US/10865978
; Publication No. US20050003416A1
; GENERAL INFORMATION:
; APPLICANT: Berlex Biosciences
; APPLICANT: Wu, Qingyu
; TITLE OF INVENTION: Novel Modified Corin Molecules Having Substitute Activation Sequ
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 53103AUST1
; CURRENT APPLICATION NUMBER: US/10/865, 978
; CURRENT FILING DATE: 2004-06-09
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 37
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Igk Secretion Signal Sequence, Corin AA 787 to 796, Enterokinase
; OTHER INFORMATION: Site, Corin AA 802 to 1042, and C-Terminal V5 and 6xHis Tags
US-10-865-978-37

Query Match 65.0%; Score 58.5; DB 16; Length 326;
Best Local Similarity 57.1%; Pred. No. 0.93;
Matches 16; Conservative 0; Mismatches 1; Indels 11; Gaps 2;

QY 1 KGEGLPIPNPLRTG 17
|||||
Db 293 KGEGLPIPNPLRTG 320

RESULT 5
US-10-467-243-26
; Sequence 26, Application US/10467243
; Publication No. US20040132971A1
; GENERAL INFORMATION:
; APPLICANT: Maxygen Holdings Ltd.
; APPLICANT: Haaning, Jesper Mortensen
; APPLICANT: Halkier, Torben
; TITLE OF INVENTION: RANK LIGAND-BINDING POLYPEPTIDES
; FILE REFERENCE: 0226wo310
; CURRENT APPLICATION NUMBER: US/10/467,243
; CURRENT FILING DATE: 2003-08-06
; PRIOR APPLICATION NUMBER: DK PA 2001 00214
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/267,843
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: DK PA 2001 00498
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/278,320
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: soluble hOPG polypeptide encoded by SEQ ID NO:25
US-10-467-243-26

Query Match 65.0%; Score 58.5; DB 16; Length 338;
Best Local Similarity 59.1%; Pred. No. 0.97;
Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

QY 1 KGEGLPIPNPLRTG 17
|||||
Db 311 KGEGLPIPNPLRTG 332

RESULT 6
US-10-467-243-8
; Sequence 8, Application US/10467243
; Publication No. US20040132971A1
; GENERAL INFORMATION:
; APPLICANT: Maxygen Holdings Ltd.
; APPLICANT: Haaning, Jesper Mortensen
; APPLICANT: Halkier, Torben
; TITLE OF INVENTION: RANK LIGAND-BINDING POLYPEPTIDES
; FILE REFERENCE: 0226wo310
; CURRENT APPLICATION NUMBER: US/10/467,243
; CURRENT FILING DATE: 2003-08-06
; PRIOR APPLICATION NUMBER: DK PA 2001 00214
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/267,843
; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: DK PA 2001 00498
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/278,320
; PRIOR FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: polypeptide encoded by SEQ ID NO:7
US-10-467-243-8

Query Match 65.0%; Score 58.5; DB 16; Length 349;
Best Local Similarity 59.1%; Pred. No. 1;
Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

Qy 1 KGELEGLPIPNPLL-----RTG 17
Db 322 RGPFGKPIPNPLGLDSTRTG 343

RESULT 7
US-10-270-223-4
; Sequence 4, Application US/10270223
; Publication No. US20030143634A1
; GENERAL INFORMATION:
; APPLICANT: Biologie A/S
; TITLE OF INVENTION: AN IMPROVED METHOD TO DETECT INTERACTIONS BETWEEN CELLULAR COMPONENTS
; TITLE OF INVENTION: INTERACT LIVING CELLS, AND TO EXTRACT QUANTITATIVE INFORMATION RE
; FILE REFERENCE: 3759-0126P
; CURRENT APPLICATION NUMBER: US/10/270,223
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Aequoria Victoria and Human
US-10-270-223-4

Query Match 65.0%; Score 58.5; DB 14; Length 368;
Best Local Similarity 59.1%; Pred. No. 1.1;
Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

Qy 1 KGELEGLPIPNPLL-----RTG 17
Db 341 RGPFGKPIPNPLGLDSTRTG 362

RESULT 8
US-10-239-652A-6
; Sequence 6, Application US/10239652A
; Publication No. US20040053234A1
; GENERAL INFORMATION:
; APPLICANT: Michael David Winther; Heidi Lynn Smith; Andre Ponton;
; APPLICANT: Roberto Justo De Antueno; Stephen John Allen
; TITLE OF INVENTION: Polynucleotides that Control Delta-6-Desaturase Genes
; TITLE OF INVENTION: and Methods for Identifying Compounds for Modulating
; TITLE OF INVENTION: Delta-6-Desaturase
; FILE REFERENCE: 42320-0010
; CURRENT APPLICATION NUMBER: US/10/239,652A
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: PCT/CA01/00398
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: CA2,301,158
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-415-232-4

Query Match 65.0%; Score 58.5; DB 15; Length 473;
Best Local Similarity 59.1%; Pred. No. 1.4;
Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

; TYPE: PRT
; ORGANISM: rat
US-10-239-652A-6

Query Match 65.0%; Score 58.5; DB 15; Length 473;
Best Local Similarity 59.1%; Pred. No. 1.4;
Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

Qy 1 KGELEGLPIPNPLL-----RTG 17
Db 446 RGPFGKPIPNPLGLDSTRTG 467

RESULT 9
US-10-239-652A-8
; Sequence 8, Application US/10239652A
; Publication No. US20040053234A1
; GENERAL INFORMATION:
; APPLICANT: Michael David Winther; Heidi Lynn Smith; Andre Ponton;
; APPLICANT: Roberto Justo De Antueno; Stephen John Allen
; TITLE OF INVENTION: Polynucleotides that Control Delta-6-Desaturase Genes
; TITLE OF INVENTION: and Methods for Identifying Compounds for Modulating
; TITLE OF INVENTION: Delta-6-Desaturase
; FILE REFERENCE: 42320-0010
; CURRENT APPLICATION NUMBER: US/10/239,652A
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: PCT/CA01/00398
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: CA2,301,158
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Human
US-10-239-652A-8

Query Match 65.0%; Score 58.5; DB 15; Length 473;
Best Local Similarity 59.1%; Pred. No. 1.4;
Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

Qy 1 KGELEGLPIPNPLL-----RTG 17
Db 446 RGPFGKPIPNPLGLDSTRTG 467

RESULT 10
US-10-415-232-4
; Sequence 4, Application US/10415232
; Publication No. US20040096435A1
; GENERAL INFORMATION:
; APPLICANT: Winther et al.
; TITLE OF INVENTION: Methods for Screening Compounds that Modulate Lipid
; TITLE OF INVENTION: Metabolism
; FILE REFERENCE: 04511.0005.NPUS01
; CURRENT APPLICATION NUMBER: US/10/415,232
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: PCT/CA01/01520
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 60/243,009
; PRIOR FILING DATE: 2000-10-26
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-415-232-4

Query Match 65.0%; Score 58.5; DB 15; Length 473;
Best Local Similarity 59.1%; Pred. No. 1.4;
Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

RESULT 15
US-10-653-496A-25
; Sequence 25, Application US/10653496A
; Publication No. US20040116672A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/10/653,496A
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/09/328,877D
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: V5HIS tagged
; OTHER INFORMATION: homo sapien
US-10-653-496A-25

Query Match 65.0%; Score 58.5; DB 16; Length 669;
Best Local Similarity 59.1%; Pred. No. 2.1;
Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

Qy 1 KGELEGPIPNPLL-----RTG 17
Db 642 RGPFGKPIPNPLLGLDSTRTG 663

Search completed: May 27, 2005, 14:15:04
Job time : 5.92964 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:28:05 ; Search time 6.45203 Seconds
(without alignments)
1019.048 Million cell updates/sec

Title: US-10-622-893A-4
Perfect score: 90
Sequence: 1 KGELEGLPNPLRTG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59.5	66.1	32	8	ADN61411 Paenibaci
2	59	65.6	146	5	AAU76975 Human tra
3	58.5	65.0	42	8	ADN40846 Synthetic
4	58.5	65.0	250	4	AAB20357 Anti-chel
5	58.5	65.0	254	8	ADQ98103 Chimeric
6	58.5	65.0	254	8	ADQ98105 Chimeric
7	58.5	65.0	254	8	ADQ98104 Chimeric
8	58.5	65.0	254	8	ADQ98107 Chimeric
9	58.5	65.0	254	8	ADQ98106 Chimeric
10	58.5	65.0	254	8	ADQ98108 Chimeric
11	58.5	65.0	338	5	ABG71834 Soluble h
12	58.5	65.0	349	5	ABG71826 Protein e
13	58.5	65.0	368	6	ABR40351 Human ami
14	58.5	65.0	374	3	AAB10639 Human VEG
15	58.5	65.0	473	4	AAE11082 C-termina
16	58.5	65.0	473	4	AAE11084 C-termina
17	58.5	65.0	473	5	AAE14740 C-termina
18	58.5	65.0	669	2	AAI22470 Human Cte
19	58.5	65.0	669	4	AAB59202 C-termina
20	56.5	62.8	253	8	ADJ62976 Novel flu
21	56.5	62.8	361	6	ABR61836 P. falcip
22	56.5	62.8	383	5	ABB07412 Amino aci
23	56.5	62.8	383	6	ABG72053 Human Hsk
24	55.5	61.7	570	8	ADR86696 Ephrin B4
25	55.5	61.7	570	8	ADR82643 Human B4E

26	55.5	61.7	1620	8	ADH71972	Adh71972 Human pro
27	55.5	61.7	1653	8	ADH71970	Adh71970 Human pro
28	51.5	57.2	28	8	ADQ48584	Adq48584 Viral vec
29	51.5	57.2	33	8	ADQ48586	Adq48586 Viral vec
30	51.5	57.2	33	8	ADQ48573	Adq48573 Viral vec
31	51.5	57.2	42	8	ADQ48581	Adq48581 Viral vec
32	51.5	57.2	42	8	ADQ48577	Adq48577 Viral vec
33	51.5	57.2	42	8	ADQ48579	Adq48579 Viral vec
34	51.5	57.2	135	8	ADJ63148	Adj63148 Novel flu
35	51.5	57.2	181	8	ADJ63042	Adj63042 Novel flu
36	51.5	57.2	196	8	ADJ63116	Adj63116 Novel flu
37	51.5	57.2	200	8	ADJ63004	Adj63004 Novel flu
38	51.5	57.2	212	8	ADJ63024	Adj63024 Novel flu
39	51.5	57.2	233	8	ADJ63128	Adj63128 Novel flu
40	51.5	57.2	236	8	ADJ63012	Adj63012 Novel flu
41	51.5	57.2	237	8	ADJ63082	Adj63082 Novel flu
42	51.5	57.2	237	8	ADJ63058	Adj63058 Novel flu
43	51.5	57.2	237	8	ADJ63144	Adj63144 Novel flu
44	51.5	57.2	237	8	ADJ63022	Adj63022 Novel flu
45	51.5	57.2	237	8	ADJ63080	Adj63080 Novel flu

ALIGNMENTS

RESULT 1
ADN61411
ID ADN61411 standard; peptide; 32 AA.
XX
AC ADN61411;
XX
DT 01-JUL-2004 (first entry)
XX
DE Paenibacillus sp. Cry1529-related V5 Ab epitope/His tag purity peptide.
XX
KW Cry; toxic; lepidopteran pest; toxin complex; insecticide; DAS1529; ORF7;
KW Cry1529; purification; V5 antibody epitope; His tag.
XX
OS Synthetic.
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Region 10..26
FT /note= "V5 antibody epitope"
XX
PN WO2004002223-A2.
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020082.
XX
PR 28-JUN-2002; 2002US-0392633P.
PR 21-JAN-2003; 2003US-0441647P.
XX
PA (DOWC) DOW AGROSCIENCES LLC.
XX
PI Bintrim SB, Bevan SA, Zhu B, Merlo DJ;
XX
DR WPI; 2004-082821/08.
XX
PT Screening a culture of Paenibacillus isolate for Cry protein or toxin
PT complex protein, useful for controlling lepidoptera, comprises
PT obtaining DNA or protein from the culture and assaying the presence of
PT the gene or protein.
XX
PS Example 11; Page 70; 220pp; English.
XX
CC The invention relates to a novel method for screening a culture of a
CC Paenibacillus isolate for a gene encoding a protein selected from a Cry
CC protein that is toxic to a lepidopteran pest and a toxin complex protein.
CC The method comprises obtaining DNA from the culture and assaying the DNA
CC for the presence of the gene or obtaining a protein produced by the
CC culture and assaying the presence of a protein that indicates the


```

XX Meares C, Corneillie T;
XX WPI; 2004-580725/56.
XX N-PSDB; ADQ98096.
XX
XX Novel mutant antibody comprising reactive site not present in wild-type
XX of antibody and antigen recognition domain that recognizes macrocyclic
XX metal chelate having four nitrogen atoms, useful for treating cancer or
XX autoimmune diseases.
XX
XX Claim 41; SEQ ID NO 80; 208pp; English.
XX
XX This invention relates to multi-functional antibodies that recognise
XX chelating agents and metal chelates, particularly macrocyclic metal
XX chelates. Specifically, it refers to an antibody that comprises a metal
XX chelate bound to an antigen recognition domain, where the metal chelate
XX has a reactive functional group of complementary reactivity to the
XX reactive site of the antibody. This reactive site is the side chain of a
XX naturally occurring amino acid e.g. the -SH group side chain of a
XX cysteine residue (not present in the wild type) which can be used to form
XX a covalent bond between the reactive site of the antibody and the
XX reactive functional group of the metal chelate. The present invention
XX describes using these antibodies for in vivo imaging where the antibody
XX comprises a targeting moiety that binds specifically to a cell via a cell
XX surface receptor or antigen thus forming a cell-mutant antibody complex.
XX On addition of the metal chelate, a cell-antibody-metal chelate complex
XX is formed that can be detected using emission tomography, magnetic
XX resonance imaging, lanthanide luminescence, gamma-emissions or single
XX photon emission tomography (SPET). As such, this method is useful for
XX treating a subject with cancer and pharmaceutical compositions exhibit
XX cytostatic and immunosuppressive activities. This polypeptide sequence is
XX a chimeric murine antibody 2D12.5 variable heavy chain protein fused to
XX the human anti-tetanus toxin antibody CH1 heavy chain constant region
XX with a V5 epitope and (His)6 tag, given in an exemplification of the
XX invention.
XX
XX Sequence 254 AA;
XX
XX Query Match 65.0%; Score 58.5; DB 8; Length 254;
XX Best Local Similarity 59.1%; Pred. No. 0.58;
XX Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;
XX
XX QY 1 KGELEGLPIPNPLL-----RTG 17
XX Db 227 RGFPEKGKPIPNLLGLDSTRTG 248
XX
XX RESULT 6
XX ADQ98105
XX ID ADQ98105 standard; protein; 254 AA.
XX AC ADQ98105;
XX
XX 21-OCT-2004 (first entry)
XX
XX Chimeric 2dVH-TTCH protein with tag and an N87D/G53C mutations SeqID 83.
XX
XX murine; mouse; chimeric; human; TTCH; tetanus toxin;
XX magnetic resonance imaging; lanthanide luminescence; gamma-emissions;
XX single photon emission tomography; SPET; cancer; cytostatic;
XX immunosuppressive; multi-functional antibody; metal chelate;
XX antigen recognition domain; in vivo imaging;
XX cell-antibody-metal chelate complex; emission tomography; V5 epitope;
XX His tag; mutant; mutein.
XX
XX Mus musculus.
XX OS Homo sapiens.
XX OS Chimeric.
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX FH Misc-difference 53
XX FT

```

```

FT Misc-difference 87 /note= "Wild type Gly substituted for Cys"
FT FT /note= "Wild type Asn substituted for Asp"
XX
XX WO2004065569-A2.
XX
XX 05-AUG-2004.
XX
XX 23-JAN-2004; 2004WO-US001808.
XX
XX 23-JAN-2003; 2003US-00350555.
XX 22-JUL-2003; 2003US-00625047.
XX 31-JUL-2003; 2003US-00631258.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Meares C, Corneillie T;
XX
XX WPI; 2004-580725/56.
XX N-PSDB; ADQ98098.
XX
XX Novel mutant antibody comprising reactive site not present in wild-type
XX of antibody and antigen recognition domain that recognizes macrocyclic
XX metal chelate having four nitrogen atoms, useful for treating cancer or
XX autoimmune diseases.
XX
XX Claim 41; SEQ ID NO 82; 208pp; English.
XX
XX This invention relates to multi-functional antibodies that recognise
XX chelating agents and metal chelates, particularly macrocyclic metal
XX chelates. Specifically, it refers to an antibody that comprises a metal
XX chelate bound to an antigen recognition domain, where the metal chelate
XX has a reactive functional group of complementary reactivity to the
XX reactive site of the antibody. This reactive site is the side chain of a
XX naturally occurring amino acid e.g. the -SH group side chain of a
XX cysteine residue (not present in the wild type) which can be used to form
XX a covalent bond between the reactive site of the antibody and the
XX reactive functional group of the metal chelate. The present invention
XX describes using these antibodies for in vivo imaging where the antibody
XX comprises a targeting moiety that binds specifically to a cell via a cell
XX surface receptor or antigen thus forming a cell-mutant antibody complex.
XX On addition of the metal chelate, a cell-antibody-metal chelate complex
XX is formed that can be detected using emission tomography, magnetic
XX resonance imaging, lanthanide luminescence, gamma-emissions or single
XX photon emission tomography (SPET). As such, this method is useful for
XX treating a subject with cancer and pharmaceutical compositions exhibit
XX cytostatic and immunosuppressive activities. This polypeptide is the
XX mutant protein sequence comprising a chimeric murine antibody 2D12.5
XX variable heavy chain protein fused to the human anti-tetanus toxin
XX antibody CH1 heavy chain constant region (2dVH-TTCH) with N87D and G53C
XX mutations and a V5 epitope with (His)6 tag, given in an exemplification
XX of the invention.
XX
XX Sequence 254 AA;
XX
XX Query Match 65.0%; Score 58.5; DB 8; Length 254;
XX Best Local Similarity 59.1%; Pred. No. 0.58;
XX Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;
XX
XX QY 1 KGELEGLPIPNPLL-----RTG 17
XX Db 227 RGFPEKGKPIPNLLGLDSTRTG 248
XX
XX RESULT 7
XX ADQ98104
XX ID ADQ98104 standard; protein; 254 AA.
XX AC ADQ98104;
XX
XX 21-OCT-2004 (first entry)
XX
XX Chimeric 2dVH-TTCH protein with tag and an N87D mutation SeqID 82.
XX

```

XX murine; mouse; chimeric; human; TTCH; tetanus toxin;
KW magnetic resonance imaging; lanthanide luminescence; gamma-emissions;
KW single photon emission tomography; SPET; cancer; cytostatic;
KW immunosuppressive; multi-functional antibody; metal chelate;
KW antigen recognition domain; in vivo imaging;
KW cell-antibody-metal chelate complex; emission tomography; V5 epitope;
KW His tag; mutant; mutein.
XX
OS Mus musculus.
OS Homo sapiens.
OS Chimeric.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 87 /note= "Wild type Asn substituted for Asp"
FT
XX
XX PN WO2004065569-A2.
XX PD 05-AUG-2004.
XX PF 23-JAN-2004; 2004WO-US001808.
XX PF 23-JAN-2003; 2003US-00350555.
XX PR 22-JUL-2003; 2003US-00625047.
XX PR 31-JUL-2003; 2003US-00631258.
XX
XX PA (REGC) UNIV CALIFORNIA.
XX PI Meares C, Corneillie T;
XX
XX DR WPI; 2004-580725/56.
XX DR N-PSDB; ADQ98097.
XX
XX PT Novel mutant antibody comprising reactive site not present in wild-type
PT of antibody and antigen recognition domain that recognizes macrocyclic
PT metal chelate having four nitrogen atoms, useful for treating cancer or
PT autoimmune diseases.
XX
XX PS Claim 41; SEQ ID NO 81; 208pp; English.
XX
XX CC This invention relates to multi-functional antibodies that recognise
CC chelating agents and metal chelates, particularly macrocyclic metal
CC chelates. Specifically, it refers to an antibody that comprises a metal
CC chelate bound to an antigen recognition domain, where the metal chelate
CC has a reactive functional group of complementary reactivity to the
CC reactive site of the antibody. This reactive site is the side chain of a
CC naturally occurring amino acid e.g. the -SH group side chain of a
CC cysteine residue (not present in the wild type) which can be used to form
CC a covalent bond between the reactive site of the antibody and the
CC reactive functional group of the metal chelate. The present invention
CC describes using these antibodies for in vivo imaging where the antibody
CC comprises a targeting moiety that binds specifically to a cell via a cell
CC surface receptor or antigen thus forming a cell-mutant antibody complex.
CC On addition of the metal chelate, a cell-antibody-metal chelate complex
CC is formed that can be detected using emission tomography, magnetic
CC resonance imaging, lanthanide luminescence, gamma-emissions or single
CC photon emission tomography (SPET). As such, this method is useful for
CC treating a subject with cancer and pharmaceutical compositions exhibit
CC cytostatic and immunosuppressive activities. This polypeptide is the
CC mutant protein sequence comprising a chimeric murine antibody 2D12.5
CC variable heavy chain protein fused to the human anti-tetanus toxin
CC antibody CH1 heavy chain constant region (2dVH-TTCH) with an N87D
CC mutation and a V5 epitope with (His)6 tag, given in an exemplification of
CC the invention.
XX
SQ Sequence 254 AA;
Query Match 65.0%; Score 58.5; DB 8; Length 254;
Best Local Similarity 59.1%; Pred.No. 0.59; Indels 5; Gaps 1;
Matches 13; Conservative 1; Mismatches 3

OY 1 KGSLEGLPIPNPLL-----RTG 17
DB 227 RGPFGKPIPNPLLGLDSTRTG 248
RESULT 8
ADQ98107
ID ADQ98107 standard; protein; 254 AA.
XX AC ADQ98107;
XX AC
XX XX 21-OCT-2004 (first entry)
XX DE Chimeric 2dVH-TTCH protein with tag and an N87D/G55C mutations SeqID 85.
XX KW murine; mouse; chimeric; human; TTCH; tetanus toxin;
KW magnetic resonance imaging; lanthanide luminescence; gamma-emissions;
KW single photon emission tomography; SPET; cancer; cytostatic;
KW immunosuppressive; multi-functional antibody; metal chelate;
KW antigen recognition domain; in vivo imaging;
KW cell-antibody-metal chelate complex; emission tomography; V5 epitope;
KW His tag; mutant; mutein.
XX
XX OS Mus musculus.
XX OS Homo sapiens.
XX OS Chimeric.
XX OS Synthetic.
XX PH Key Location/Qualifiers
FT Misc-difference 55 /note= "Wild type Gly substituted for Cys"
FT Misc-difference 87 /note= "Wild type Asn substituted for Asp"
FT
XX XX WO2004065569-A2.
XX XX 05-AUG-2004.
XX XX 23-JAN-2004; 2004WO-US001808.
XX XX 23-JAN-2003; 2003US-00350555.
XX XX 22-JUL-2003; 2003US-00625047.
XX XX 31-JUL-2003; 2003US-00631258.
XX
XX PA (REGC) UNIV CALIFORNIA.
XX PI Meares C, Corneillie T;
XX
XX DR WPI; 2004-580725/56.
XX DR N-PSDB; ADQ98100.
XX
XX PT Novel mutant antibody comprising reactive site not present in wild-type
PT of antibody and antigen recognition domain that recognizes macrocyclic
PT metal chelate having four nitrogen atoms, useful for treating cancer or
PT autoimmune diseases.
XX
XX PS Claim 41; SEQ ID NO 84; 208pp; English.
XX
XX CC This invention relates to multi-functional antibodies that recognise
CC chelating agents and metal chelates, particularly macrocyclic metal
CC chelates. Specifically, it refers to an antibody that comprises a metal
CC chelate bound to an antigen recognition domain, where the metal chelate
CC has a reactive functional group of complementary reactivity to the
CC reactive site of the antibody. This reactive site is the side chain of a
CC naturally occurring amino acid e.g. the -SH group side chain of a
CC cysteine residue (not present in the wild type) which can be used to form
CC a covalent bond between the reactive site of the antibody and the
CC reactive functional group of the metal chelate. The present invention
CC describes using these antibodies for in vivo imaging where the antibody
CC comprises a targeting moiety that binds specifically to a cell via a cell
CC surface receptor or antigen thus forming a cell-mutant antibody complex.
CC On addition of the metal chelate, a cell-antibody-metal chelate complex
CC is formed that can be detected using emission tomography, magnetic

CC resonance imaging, lanthanide luminescence, gamma-emissions or single
 CC photon emission tomography (SPET). As such, this method is useful for
 CC treating a subject with cancer and pharmaceutical compositions exhibit
 CC cytostatic and immunosuppressive activities. This polypeptide is the
 CC mutant protein sequence comprising a chimeric murine antibody 2D12.5
 CC variable heavy chain protein fused to the human anti-tetanus toxin
 CC antibody CH1 heavy chain constant region (2dVH-TTCH) with N87D and G54C
 CC mutations and a V5 epitope with (His)6 tag, given in an exemplification
 CC of the invention.
 XX
 SQ Sequence 254 AA;

Query Match 65.0%; Score 58.5; DB 8; Length 254;
 Best Local Similarity 59.1%; Pred. NO. 0.58;
 Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

QY 1 KGELEGPIPNPLL-----RTG 17
 Db 227 RGPFEKPIPNPLGLDSTRTG 248

RESULT 9
 ADQ98106
 ID ADQ98106 standard; protein; 254 AA.
 XX
 AC ADQ98106;

XX
 DT 21-OCT-2004 (first entry)

XX Chimeric 2dVH-TTCH protein with tag and an N87D/G54C mutations SeqID 84.
 DE
 KW murine; mouse; chimeric; human; TTCH; tetanus toxin;
 KW magnetic resonance imaging; lanthanide luminescence; gamma-emissions;
 KW single photon emission tomography; SPET; cancer; cytostatic;
 KW immunosuppressive; multi-functional antibody; metal chelate;
 KW antigen recognition domain; in vivo imaging;
 KW cell-antibody-metal chelate complex; emission tomography; V5 epitope;
 KW His tag; mutant; mutein.

XX Mus musculus.
 OS Homo sapiens.
 OS Chimeric.
 OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 54 /note= "Wild type Gly substituted for Cys"

FT Misc-difference 87 /note= "Wild type Asn substituted for Asp"

FT

XX WO2004065569-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC This invention relates to multi-functional antibodies that recognize
 CC chelating agents and metal chelates, particularly macrocyclic metal
 CC chelates. Specifically, it refers to an antibody that comprises a metal
 CC chelate bound to an antigen recognition domain, where the metal chelate
 CC has a reactive functional group of complementary reactivity to the
 CC reactive site of the antibody. This reactive site is the side chain of a
 CC naturally occurring amino acid e.g. the -SH group side chain of a
 CC cysteine residue (not present in the wild type) which can be used to form
 CC a covalent bond between the reactive site of the antibody and the
 CC reactive functional group of the metal chelate. The present invention
 CC describes using these antibodies for in vivo imaging where the antibody
 CC comprises a targeting moiety that binds specifically to a cell via a cell
 CC surface receptor or antigen thus forming a cell-mutant antibody complex.
 CC On addition of the metal chelate, a cell-antibody-metal chelate complex
 CC is formed that can be detected using emission tomography, magnetic
 CC resonance imaging, lanthanide luminescence, gamma-emissions or single
 CC photon emission tomography (SPET). As such, this method is useful for
 CC treating a subject with cancer and pharmaceutical compositions exhibit
 CC cytostatic and immunosuppressive activities. This polypeptide is the
 CC mutant protein sequence comprising a chimeric murine antibody 2D12.5
 CC variable heavy chain protein fused to the human anti-tetanus toxin
 CC antibody CH1 heavy chain constant region (2dVH-TTCH) with N87D and G54C
 CC mutations and a V5 epitope with (His)6 tag, given in an exemplification
 CC of the invention.
 XX
 SQ Sequence 254 AA;

Query Match 65.0%; Score 58.5; DB 8; Length 254;
 Best Local Similarity 59.1%; Pred. NO. 0.58;
 Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

QY 1 KGELEGPIPNPLL-----RTG 17
 Db 227 RGPFEKPIPNPLGLDSTRTG 248

RESULT 10

ADQ98108

ID ADQ98108 standard; protein; 254 AA.

XX

AC ADQ98108;

XX

XX 21-OCT-2004 (first entry)

XX Chimeric murine 2D12.5 VH antibody fused to human TTCH protein with tag.

XX murine; mouse; chimeric; human; TTCH; tetanus toxin;

XX magnetic resonance imaging; lanthanide luminescence; gamma-emissions;

XX single photon emission tomography; SPET; cancer; cytostatic;

XX immunosuppressive; multi-functional antibody; metal chelate;

XX antigen recognition domain; in vivo imaging;

XX cell-antibody-metal chelate complex; emission tomography; V5 epitope;

XX His tag.

XX

XX Mus musculus.

OS Homo sapiens.

OS Chimeric.

OS Synthetic.

XX

XX WO2004065569-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX


```

XX 09-FEB-2001; 2001DK-00000214.
PR 09-FEB-2001; 2001US-0267843P.
PR 23-MAR-2001; 2001DK-00000498.
PR 23-MAR-2001; 2001US-0278320P.
XX (MAXY-) MAXYGEN HOLDINGS LTD.
PA Haaning JM, Halkier T;
XX WPI; 2002-691592/74.
XX N-PSDB; ABS54847.
XX Novel human receptor activator of NFkappaB (hRANK) or human
PT osteoprotegerin (hOPG) variant polypeptides which bind to RANK ligand
PT (RANKL) with equivalent binding affinity as hRANK or hOPG, useful for
PT treating osteoporosis.
XX Disclosure; Fig 6; 129pp; English.
XX This invention relates to a novel polypeptide having an amino acid
CC sequence that is different from and is at least about 70% identical to
CC the amino acid sequence of human receptor activator of NFkappaB (hRANK)
CC or human osteoprotegerin (hOPG), and which has a binding affinity to RANK
CC ligand (RANKL) that is at least as high as the binding affinity of hRANK
CC or hOPG to RANKL, as determined by functional competition assay. The
CC protein of the invention may have osteopathic activity and may act as a
CC RANKL-mediated osteoclastogenesis or RANKL-mediated osteoclast activity
CC inhibitor. The nucleotide sequence shown in the invention may be used in
CC gene therapy. The protein of the invention or fusion proteins comprising
CC this protein are useful as a pharmaceutical, and in the preparation of a
CC medicament for treating or preventing osteoporosis, or other bone
CC diseases or diseases associated with binding of RANKL to the RANK
CC receptor. A host cell containing a vector expressing the protein is
CC useful for producing a polypeptide having binding affinity to RANKL,
CC where the polypeptide comprises at least one N- or O-glycosylation site
CC and the host cell is a eukaryotic host cell capable of in vivo
CC glycosylation, and/or the polypeptide is subjected to conjugation to a
CC non-polypeptide moiety in vitro. The protein of the invention has
CC increased functional in vivo half-life and/or serum half-life compared to
CC hRANK or hOPG and has an improved binding affinity to RANKL compared to
CC the binding affinity of hRANK or hOPG to RANKL, as determined by a
CC functional competition assay. The present sequence represents a human
CC RANKL protein-HIS tag fusion protein encoded by the pYhRANKb expression
CC vector used in the method of the invention
XX Sequence 349 AA;
SQ
Query Match 65.0%; Score 58.5; DB 5; Length 349;
Best Local Similarity 59.1%; Pred. No. 0.81;
Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;
QY 1 KGELEGPIPNPLL-----RTG 17
DB 322 RGFPEKPIPNPLGLDSTRTG 343
RESULT 13
ABR40351
ID ABR40351 standard; protein; 368 AA.
XX ABR40351;
XX OS Homo sapiens.
XX AC ABR40351;
XX DT 08-JUL-2003 (first entry)
XX DE Human amino acid sequence SEQ ID NO: 4.
XX KW Human; heterologous conjugate; intracellular protein.
XX OS Homo sapiens.
XX OS Aequoria victoria.
XX KW WO2003029827-A2.
XX PN
XX
XX 10-APR-2003.
XX 01-OCT-2002; 2002WO-DK000651.
XX 01-OCT-2001; 2001DK-00001433.
XX 11-OCT-2001; 2001US-0328896P.
XX (BIOL-) BIOLMAGE AS.
XX Terry BR, Nielsen SJ;
XX WPI; 2003-430211/40.
XX N-PSDB; ACC72603.
XX Novel cell for identifying modulators of protein interaction, contains a
PT first conjugate comprising anchor protein, second conjugate having type B
PT interactor protein and third conjugate with detectable group.
XX Disclosure; Page 110; 118pp; English.
XX The invention relates to a novel cell, comprising three heterologous
CC conjugates (HC), a first HC (HC1) comprising an anchor protein that
CC specifically binds to an internal structure within the cell conjugated to
CC an interactor protein (IP) of type A, a second HC (HC2) comprising IP of
CC type B conjugated to a first protein of interest, and a third HC (HC3)
CC comprising a second protein of interest conjugated to detectable group.
CC The cell is useful for detecting if a compound disrupts or induces the
CC interaction between two intracellular proteins. The cell is also useful
CC for screening compounds that modulate the interaction between two
CC intracellular proteins. The present sequence is used in the
CC exemplification of the invention
XX Sequence 368 AA;
SQ
Query Match 65.0%; Score 58.5; DB 6; Length 368;
Best Local Similarity 59.1%; Pred. No. 0.86;
Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;
QY 1 KGELEGPIPNPLL-----RTG 17
DB 341 RGFPEKPIPNPLGLDSTRTG 362
RESULT 14
AAB10639
ID AAB10639 standard; protein; 374 AA.
XX AAB10639;
XX AC AAB10639;
XX DT 19-JAN-2001 (first entry)
XX DE Human VEGF-X protein for expression in mammalian systems.
XX KW VEGF-X; vascular endothelial growth factor; human; vulnerary; cytostatic;
XX anti-rheumatic; antiarthritic; antipsoriatic; antidiabetic; treatment;
XX angiogenesis regulator; vascularization regulator; cancer; psoriasis;
XX rheumatoid arthritis; diabetic retinopathy; blood vessel; organ repair;
XX tissue regeneration; tissue repair; wound; dermal ulcer; pressure sore;
XX venous sore; diabetic ulcer; burns; skin graft growth.
XX OS Homo sapiens.
XX AC AAB10639;
XX DT 29-JUN-2000.
XX DE 21-DEC-1999; 99WO-US030503.
XX KW 22-DEC-1998; 98GB-00028377.
XX PR 18-MAR-1999; 99US-0124967P.
XX PR 08-NOV-1999; 99US-0164131P.
XX PN

```

PA	(JANC) JANSSEN PHARM NV.
XX	
PI	Gordon RD, Sprengel JJ, Yon JR, Dijkmans JJH, Gosiewska A;
PI	Dhanaraj SN, Xu J;
XX	
DR	WPI; 2000-442669/38.
DR	N-PSDB; AAA71983.
XX	
PT	New vascular endothelial growth factor protein, useful for treating or
PT	preventing diseases associated with inappropriate angiogenesis activity
PT	such as cancer, rheumatoid arthritis, psoriasis and wounds.
XX	
PS	Disclosure; Fig 19; 127pp; English.
XX	
CC	This invention describes a novel vascular endothelial growth factor-X
CC	(VEGF-X) protein (Ia) and its encoding polynucleotide (IIa) which has
CC	vulnerary, cytostatic, antirheumatic, antiarthritic, antipsoriatic and
CC	antidiabetic activity and acts as an angiogenesis and vasculization
CC	regulator. An antisense molecule of the invention is useful for treating
CC	or preventing cancer, rheumatoid arthritis, psoriasis and diabetic
CC	retinopathy by inhibiting angiogenic activity or inappropriate
CC	vascularization including formation and proliferation of new blood
CC	vessels, growth and development of tissues, tissue regeneration and organ
CC	and tissue repair in a subject. The products of the invention are useful
CC	for preparing medicaments for treating wounds such as dermal ulcers,
CC	pressure sores, venous sores, diabetic ulcers and burns and to promote
CC	skin graft growth, tissue repair, proliferation of new blood vessels,
CC	tissue regeneration and organ repair by promoting angiogenic activity or
CC	vascularization. This sequence represents a human VEGF-X protein which
CC	can be expressed in mammalian systems and which is described in the
CC	method of the invention
XX	
SQ	Sequence 374 AA;
	Query Match 65.0%; Score 58.5; DB 3; Length 374;
	Best Local Similarity 59.1%; Pred.No. 0.87;
	Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;
QY	1 KGELEGLIPNPILL-----RTG 17
	:
Db	347 RGPFEKPIPPLGLDSTRTG 368
RESULT 15	
AAE11082	
ID	AAE11082 standard; protein; 473 AA.
AC	AAE11082;
XX	
DT	18-DEC-2001 (first entry)
XX	
DE	C-terminal tagged rat delta-6-desaturase (rD6D-1).
XX	
KW	Delta-6-desaturase gene; D6D; lipid metabolism disorder; atopic eczema;
KW	mastalgia; rheumatoid arthritis; Sjogren's syndrome; viral infection;
KW	gastrointestinal disorder; post viral fatigue; pre-menstrual syndrome;
KW	endometriosis; cystic fibrosis; alcoholism; Alzheimer's disease;
KW	cardiovascular disease; Crohn's disease; congenital liver disease;
KW	schizophrenia; diabetic neuropathy; nephropathy; retinopathy; cancer;
KW	arterial hypertension; atherosclerosis; chronic inflammatory disorder;
KW	autoimmune disorder; hypercholesterolaemia; atopic disorder; rD6D-1;
XX	gene therapy; rat.
XX	
OS	Rattus sp.
XX	
FH	Key Location/Qualifiers
FT	Region 451..464
FT	/note= "V5 tag"
FT	Region 468..473
FT	/note= "6xHis tag"
XX	
PX	WO200170993-A2.
PN	
XX	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:35:35 ; Search time 1.3049 Seconds
(without alignments)
972.511 Million cell updates/sec

Title: US-10-622-893A-4
Perfect score: 90
Sequence: 1 KGELEGPIPNPLRTG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	65.6	146	4	US-09-923-304-9
2	58.5	65.0	374	4	US-09-468-647A-118
3	58.5	65.0	669	4	US-09-213-888-25
4	58.5	65.0	669	4	US-09-328-877D-25
5	56.5	62.8	383	4	US-09-883-096-5
6	51.5	57.2	635	4	US-09-545-814-32
7	45	50.0	758	1	US-07-756-250-16
8	45	50.0	766	4	US-09-949-016-10058
9	44	48.9	410	4	US-09-489-039A-10689
10	44	48.9	1013	2	US-08-866-650-3
11	44	48.9	1013	2	US-09-021-287-3
12	44	48.9	1013	3	US-09-240-473-3
13	43	47.8	14	4	US-09-352-171-14
14	43	47.8	622	4	US-09-949-016-9104
15	42	46.7	271	4	US-09-270-767-60778
16	42	46.7	313	4	US-09-902-540-13536
17	42	46.7	348	4	US-09-489-039A-8827
18	42	46.7	393	4	US-09-252-991A-20362
19	42	46.7	473	3	US-08-857-076-99
20	42	46.7	591	3	US-08-951-408-4
21	42	46.7	591	3	US-09-432-473-4
22	42	46.7	778	4	US-10-148-806-36
23	42	46.7	1013	2	US-08-866-650-5
24	42	46.7	1013	2	US-09-021-287-5
25	42	46.7	1013	3	US-08-991-408-2
26	42	46.7	1013	3	US-09-240-473-5
27	42	46.7	1013	3	US-09-432-473-2

```

28 42 46.7 1013 4 US-09-285-385C-20 Sequence 20, Appl
29 42 46.7 1013 4 US-09-902-540-15897 Sequence 15897, A
30 41 45.6 232 4 US-09-134-000C-4038 Sequence 4038, Ap
31 41 45.6 429 4 US-09-902-540-12898 Sequence 12898, A
32 41 45.6 430 4 US-10-132-920B-2 Sequence 2, Appli
33 41 45.6 430 4 US-09-700-696C-2 Sequence 2, Appli
34 41 45.6 435 4 US-09-270-767-45936 Sequence 45936, A
35 41 45.6 509 4 US-09-794-422-34 Sequence 34, Appl
36 41 45.6 514 4 US-09-232-858B-14 Sequence 14, Appl
37 41 45.6 525 4 US-10-132-920B-27 Sequence 27, Appl
38 41 45.6 525 4 US-09-794-422-6 Sequence 6, Appli
39 41 45.6 540 4 US-09-794-422-46 Sequence 46, Appli
40 41 45.6 556 4 US-09-794-422-8 Sequence 8, Appli
41 41 45.6 559 4 US-10-116-370-2 Sequence 2, Appli
42 41 45.6 567 4 US-09-949-016-10952 Sequence 10952, A
43 41 45.6 706 4 US-09-248-796A-18837 Sequence 18837, A
44 41 45.6 772 4 US-10-148-806-37 Sequence 37, Appl
45 41 45.6 1509 4 US-09-677-046A-2 Sequence 2, Appli

```

ALIGNMENTS

```

RESULT 1
US-09-923-304-9
; Sequence 9, Application US/09923304
; Patent No. 6797471
; GENERAL INFORMATION:
; APPLICANT: KATZ, RUTH
; APPLICANT: JIANG, FENG
; TITLE OF INVENTION: DETECTION AND DIAGNOSIS OF SMOKING RELATED CANCERS
; FILE REFERENCE: UTSC:6580S
; CURRENT APPLICATION NUMBER: US/09/923.304
; CURRENT FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-923-304-9

```

```

Query Match 65.6%; Score 59; DB 4; Length 146;
Best Local Similarity 59.3%; Pred. No. 0.06;
Matches 16; Conservative 0; Mismatches 1; Indels 10; Gaps 2;

```

```

Qy 1 KGEI-----EGLPIPNPL-----RTG 17
    |||||
Db 114 KGEIHRHPEFGKPIPNPLGLDSTRTG 140

```

```

RESULT 2
US-09-468-647A-118
; Sequence 118, Application US/09468647A
; Patent No. 6783953
; GENERAL INFORMATION:
; APPLICANT: Gordon, Robert D
; APPLICANT: Sprengel, Jorg J
; APPLICANT: Yon, Jeffery R
; APPLICANT: Dijkmans, Josiena J.H.
; APPLICANT: Gosciewska, Anna
; APPLICANT: Dhanaraj, Sri devi N
; APPLICANT: Xu, Jean
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-X
; FILE REFERENCE: B0192.70011US00
; CURRENT APPLICATION NUMBER: US/09/468.647A
; CURRENT FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: GB 9828377.3
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 60/124,967

```

```
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 60/164,131
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 118
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-468-647A-118

Query Match          65.0%; Score 58.5; DB 4; Length 374;
Best Local Similarity 59.1%; Pred. No. 0.2;
Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

QY 1 KGEGLPIPNPLL-----RTG 17
Db 347 RGFEGKPIPNPLGLDSTRTG 368

RESULT 3
US-09-213-888-25
; Sequence 25, Application US/09213888A
; Patent No. 6638731
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; TITLE OF INVENTION: Encode Them
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213,888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VSHIS tagged
; OTHER INFORMATION: homo sapien
US-09-213-888-25

Query Match          65.0%; Score 58.5; DB 4; Length 669;
Best Local Similarity 59.1%; Pred. No. 0.39;
Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

QY 1 KGEGLPIPNPLL-----RTG 17
Db 642 RGFEGKPIPNPLGLDSTRTG 663

RESULT 4
US-09-328-877D-25
; Sequence 25, Application US/09328877D
; Patent No. 6730778
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; TITLE OF INVENTION: Encode Them
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328,877D
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Tagged Ctenocephalides felis

; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 60/164,131
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 118
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-468-647A-118

Query Match          65.0%; Score 58.5; DB 4; Length 669;
Best Local Similarity 59.1%; Pred. No. 0.39;
Matches 13; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

QY 1 KGEGLPIPNPLL-----RTG 17
Db 642 RGFEGKPIPNPLGLDSTRTG 663

RESULT 5
US-09-883-096-5
; Sequence 5, Application US/09883096
; Patent No. 6680369
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Craven, Andrew
; APPLICANT: Yu, Ming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
; FILE REFERENCE: 020552-001410US
; CURRENT APPLICATION NUMBER: US/09/883,096
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 09/594,655
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HsKip3a
; OTHER INFORMATION: fragment
; OTHER INFORMATION: Amino acid sequence of HsKip3a fragment used in
; OTHER INFORMATION: the ATPase assay (Figure 4).
US-09-883-096-5

Query Match          62.8%; Score 56.5; DB 4; Length 383;
Best Local Similarity 65.0%; Pred. No. 0.44;
Matches 13; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

QY 3 ELEGLPINPLL-----RTG 17
Db 358 KLEGKPIPNPLGLDSTRTG 377

RESULT 6
US-09-545-814-32
; Sequence 32, Application US/09545814
; Patent No. 6416977
; GENERAL INFORMATION:
; APPLICANT: Becher, Anna M.
; TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: FC-5-C1
; CURRENT APPLICATION NUMBER: US/09/545,814
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,833
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Tagged Ctenocephalides felis
```

US-09-545-814-32

Query Match 57.2%; Score 51.5; DB 4; Length 635;
Best Local Similarity 66.7%; Pred. No. 4.8;
Matches 12; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Qy 5 EGLPIPNPL--RTG 17
Db 612 EGKPIPNPLGLDSTRIG 629

RESULT 7

US-07-756-250-16
; Sequence 16, Application US/07756250

; Patent No. 5268275

; GENERAL INFORMATION:

; APPLICANT: Wu, Sheue-Mei

; APPLICANT: Stafford, Darrel W.

; TITLE OF INVENTION: Vitamin K-Dependent Carboxylase

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Kenneth D. Sibley; Bell, Seltzer, Park and

; ADDRESSEE: Gibson

; STREET: Post Office Drawer 34009

; CITY: Charlotte

; STATE: No. 5268275th Carolina

; COUNTRY: U.S.A.

; ZIP: 28234

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07756,250

; FILING DATE: 19910909

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 07/697,427

; FILING DATE: 08-MAY-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Sibley, Kenneth D.

; REGISTRATION NUMBER: 31,665

; REFERENCE/DOCKET NUMBER: 5470-34

; TELEPHONE: 919-881-3140

; TELEFAX: 919-881-3175

; TELEX: 575102

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 758 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-07-756-250-16

Query Match 50.0%; Score 45; DB 1; Length 758;

Best Local Similarity 50.0%; Pred. No. 63;

Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KGELEGLPIPNPLRT 16

Db 643 EGEVKGPEPTPLVQT 658

RESULT 8

US-09-949-016-10058

; Sequence 10058, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 10058

; LENGTH: 766

; TYPE: PRT

; ORGANISM: Human

US-09-949-016-10058

Query Match 50.0%; Score 45; DB 4; Length 766;

Best Local Similarity 50.0%; Pred. No. 64;

Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KGELEGLPIPNPLRT 16

Db 651 EGEVKGPEPTPLVQT 666

RESULT 9

US-09-489-039A-10689

; Sequence 10689, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 10689

; LENGTH: 410

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-10689

Query Match 48.9%; Score 44; DB 4; Length 410;

Best Local Similarity 58.3%; Pred. No. 46;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 6 GLPIPNPLRTG 17

Db 323 GLPLPQPTVETG 334

RESULT 10

US-08-866-650-3

; Sequence 3, Application US/08866650

; Patent No. 5939321

; GENERAL INFORMATION:

; APPLICANT: Greenspan, Daniel S

; APPLICANT: Takahara, Kazuhiko

; APPLICANT: Hoffman, Guy G

; TITLE OF INVENTION: Mammalian Tolloid-Like Protein

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Quarles & Brady

; STREET: 1 South Pinckney Street

; CITY: Madison

; STATE: WI

; COUNTRY: US

; ZIP: 53703

; COMPUTER READABLE FORM:

```
/ ; MEDIUM TYPE: Floppy disk
/ ; COMPUTER: IBM PC compatible
/ ; OPERATING SYSTEM: PC-DOS/MS-DOS
/ ; SOFTWARE: PatentIn Release #1.0, Version #1.30
/ ; CURRENT APPLICATION DATA:
/ ; APPLICATION NUMBER: US/08/866,650
/ ; FILING DATE:
/ ; CLASSIFICATION: 514
/ ; ATTORNEY/AGENT INFORMATION:
/ ; NAME: Berson, Bennett J
/ ; REGISTRATION NUMBER: 37094
/ ; REFERENCE/DOCKET NUMBER: 960296.93839
/ ; TELEPHONE: 608-251-5000
/ ; TELEFAX: 608-251-9166
/ ; INFORMATION FOR SEQ ID NO: 3:
/ ; LENGTH: 1013 amino acids
/ ; TYPE: amino acid
/ ; TOPOLOGY: linear
/ ; MOLECULE TYPE: protein
/ ; US-08-866-650-3
/ ;
/ ; Query Match 48.9%; Score 44; DB 2; Length 1013;
/ ; Best Local Similarity 56.2%; Pred. No. 1.2e+02;
/ ; Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
/ ;
/ ; QY 2 GELEGLPIPNPLRTG 17
/ ; Db 846 GRGCGSKIPDPLMATG 861
/ ;
/ ; RESULT 11
/ ; US-09-021-287-3
/ ; Sequence 3, Application US/09021287
/ ; Patent No. 5981717
/ ; GENERAL INFORMATION:
/ ; APPLICANT: Greenspan, Daniel S
/ ; APPLICANT: Takahara, Kazuhiko
/ ; APPLICANT: Hoffman, Guy G
/ ; TITLE OF INVENTION: Mammalian Tolloid-Like Protein
/ ; NUMBER OF SEQUENCES: 13
/ ; CORRESPONDENCE ADDRESS:
/ ; ADDRESSEE: Quarles & Brady
/ ; STREET: 1 South Pinckney Street
/ ; CITY: Madison
/ ; STATE: WI
/ ; COUNTRY: US
/ ; ZIP: 53703
/ ; COMPUTER READABLE FORM:
/ ; MEDIUM TYPE: Floppy disk
/ ; COMPUTER: IBM PC compatible
/ ; OPERATING SYSTEM: PC-DOS/MS-DOS
/ ; SOFTWARE: PatentIn Release #1.0, Version #1.30
/ ; CURRENT APPLICATION DATA:
/ ; APPLICATION NUMBER: US/09/021,287
/ ; FILING DATE:
/ ; CLASSIFICATION:
/ ; ATTORNEY/AGENT INFORMATION:
/ ; NAME: Berson, Bennett J
/ ; REGISTRATION NUMBER: 37094
/ ; REFERENCE/DOCKET NUMBER: 960296.93839
/ ; TELEPHONE: 608-251-5000
/ ; TELEFAX: 608-251-9166
/ ; INFORMATION FOR SEQ ID NO: 3:
/ ; LENGTH: 1013 amino acids
/ ; TYPE: amino acid
/ ; TOPOLOGY: linear
/ ; MOLECULE TYPE: protein
/ ; US-09-021-287-3
/ ;
/ ; Query Match 48.9%; Score 44; DB 2; Length 1013;
/ ; Best Local Similarity 56.2%; Pred. No. 1.2e+02;
/ ; Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
/ ;
/ ; QY 2 GELEGLPIPNPLRTG 17
/ ; Db 846 GRGCGSKIPDPLMATG 861
/ ;
/ ; RESULT 12
/ ; US-09-240-473-3
/ ; Sequence 3, Application US/09240473
/ ; Patent No. 6297011
/ ; GENERAL INFORMATION:
/ ; APPLICANT: Greenspan, Daniel S
/ ; APPLICANT: Takahara, Kazuhiko
/ ; APPLICANT: Hoffman, Guy G
/ ; TITLE OF INVENTION: Mammalian Tolloid-Like Protein
/ ; NUMBER OF SEQUENCES: 13
/ ; CORRESPONDENCE ADDRESS:
/ ; ADDRESSEE: Quarles & Brady
/ ; STREET: 1 South Pinckney Street
/ ; CITY: Madison
/ ; STATE: WI
/ ; COUNTRY: US
/ ; ZIP: 53703
/ ; COMPUTER READABLE FORM:
/ ; MEDIUM TYPE: Floppy disk
/ ; COMPUTER: IBM PC compatible
/ ; OPERATING SYSTEM: PC-DOS/MS-DOS
/ ; SOFTWARE: PatentIn Release #1.0, Version #1.30
/ ; CURRENT APPLICATION DATA:
/ ; APPLICATION NUMBER: US/09/240,473
/ ; FILING DATE:
/ ; CLASSIFICATION:
/ ; ATTORNEY/AGENT INFORMATION:
/ ; NAME: Berson, Bennett J
/ ; REGISTRATION NUMBER: 37094
/ ; REFERENCE/DOCKET NUMBER: 960296.93839
/ ; TELEPHONE: 608-251-5000
/ ; TELEFAX: 608-251-9166
/ ; INFORMATION FOR SEQ ID NO: 3:
/ ; LENGTH: 1013 amino acids
/ ; TYPE: amino acid
/ ; TOPOLOGY: linear
/ ; MOLECULE TYPE: protein
/ ; US-09-240-473-3
/ ;
/ ; Query Match 48.9%; Score 44; DB 3; Length 1013;
/ ; Best Local Similarity 56.2%; Pred. No. 1.2e+02;
/ ; Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
/ ;
/ ; QY 2 GELEGLPIPNPLRTG 17
/ ; Db 846 GRGCGSKIPDPLMATG 861
/ ;
/ ; RESULT 13
/ ; US-09-352-171-14
/ ; Sequence 14, Application US/09352171
/ ; Patent No. 6759206
/ ; GENERAL INFORMATION:
/ ; APPLICANT: Rubin, Richard A
/ ; APPLICANT: Conway, Bruce
/ ; APPLICANT: Giuliano, Kenneth A
/ ; APPLICANT: Gough, Albert H
/ ; APPLICANT: Dunlay, R. Terry
/ ; TITLE OF INVENTION: A System for Cell Based Screening
```

/

; FILE REFERENCE: 97, 022-D1
; CURRENT APPLICATION NUMBER: US/09/352,171
; CURRENT FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: epitope tag
US-09-352-171-14

Query Match 47.8%; Score 43; DB 4; Length 14;
Best Local Similarity 88.9%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 GLPIPNPLL 14
| | | | | | | |
Db 1 GKPIPNPLL 9

RESULT 14
US-09-949-016-9104
; Sequence 9104, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9104
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9104

Query Match 47.8%; Score 43; DB 4; Length 622;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 GLPIPNPLLRTG 17
| | | | | | | | | | | | | | | | | |
Db 14 GYPLGFLAQTG 25

RESULT 15
US-09-270-767-60778
; Sequence 60778, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60778
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-60778

Query Match 46.7%; Score 42; DB 4; Length 271;
Best Local Similarity 57.1%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 2 GELEGLPIPNPLL 15
| | | | | | | | | | | | | | | | | |
Db 229 GEIDGLTIPKNLLQ 242

Search completed: May 27, 2005, 13:50:00
Job time : 2.3049 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:28:55 ; Search time 3.73348 Seconds
(without alignments)
2331.700 Million cell updates/sec

Title: US-10-622-893A-4
Perfect score: 90
Sequence: 1 KGEGLPIPNPLRTG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	55.6	131	2 QBJW23	QBJW23 human adeno
2	50	55.6	131	2 QBB8S5	QBB8S5 human adeno
3	50	55.6	529	1 FTSK HAEIN	P45264 haemophilus
4	48	53.3	131	2 Q67729	Q67729 human adeno
5	48	53.3	154	2 QBP406	QBP406 xanthomonas
6	47	52.2	147	2 Q34HT5	Q34HT5 oryza sativ
7	47	52.2	147	2 Q7XG14	Q7XG14 oryza sativ
8	47	52.2	358	1 CBBA RAT	P05554 rattus norv
9	47	52.2	522	2 Q7QGZ3	Q7QGZ3 anopheles g
10	47	52.2	529	1 Y243 ARCFU	O29996 archaeoglob
11	47	52.2	579	2 Q94HML	Q94HML oryza sativ
12	47	52.2	579	2 Q7XG15	Q7XG15 oryza sativ
13	47	52.2	2039	2 Q7S3G9	Q7S3G9 neurospora
14	46	51.1	202	2 Q6MND1	Q6MND1 bdellovibri
15	46	51.1	255	2 Q8UG13	Q8UG13 agrobacteri
16	46	51.1	310	2 Q9A5C1	Q9A5C1 bacteroides
17	46	51.1	1494	2 Q6CG13	Q6CG13 yarrowia li
18	45	50.0	166	2 QBP9X3	QBP9X3 rattus norv
19	45	50.0	261	2 QBRREU8	QBRREU8 fusobacteri
20	45	50.0	263	2 Q7P3V5	Q7P3V5 fusobacteri
21	45	50.0	299	2 Q6GL93	Q6GL93 xenopus tro
22	45	50.0	372	2 Q8FZL8	Q8FZL8 brucella su
23	45	50.0	378	2 Q8YI97	Q8YI97 brucella me
24	45	50.0	560	2 Q9YIV7	Q9YIV7 rosophila
25	45	50.0	645	2 Q88RKO	Q88RKO pseudomonas
26	45	50.0	757	1 VKGC MOUSE	Q9GYC7 mus musculu
27	45	50.0	758	1 VKGC BOVIN	Q07175 bos taurus
28	45	50.0	758	1 VKGC HUMAN	P38435 homo sapien
29	45	50.0	758	1 VKGC RAT	Q88496 rattus norv
30	45	50.0	758	2 Q9GL59	Q9GL59 ovis arie
31	45	50.0	758	2 Q9MYT3	Q9MYT3 delphinapte

32	45	50.0	930	1 FTSK_PASMU	Q9CPL3 pasteurella
33	45	50.0	1138	2 Q6BZA3	Q6BZA3 debaryomyce
34	45	50.0	1371	2 P73337	P73337 synchocyst
35	44.5	49.4	356	2 Q7UYG2	Q7UYG2 rhodospirell
36	44	48.9	103	1 HEX9 ADECT	P14268 canine aden
37	44	48.9	103	2 Q7M6A0	Q7M6A0 canine aden
38	44	48.9	145	2 Q67106	Q67106 aquifex seo
39	44	48.9	257	1 TRMD_BORBR	Q7U384 bordetella
40	44	48.9	257	1 TRMD_BORPA	Q7U374 bordetella
41	44	48.9	257	1 TRMD_BORPE	Q7U361 bordetella
42	44	48.9	266	2 Q54360	Q54360 streptomyce
43	44	48.9	275	2 Q67SK7	Q67SK7 symbiobacte
44	44	48.9	339	2 Q8TQC9	Q8TQC9 methanosarc
45	44	48.9	482	2 Q8PA76	Q8PA76 xanthomonas

ALIGNMENTS

RESULT 1

Q8JW23 ID Q8JW23 PRELIMINARY; PRT; 131 AA.

AC Q8JW23; DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

DE E3 14.9kDa.

OS Human adenovirus type 34.

OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.

OX NCBI_TaxID=10548;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Compton;

RX PubMed=14693847;

RA Adhikary A.K., Inada T., Numaga J., Suzuki E., Uehijima H., Banik U., Mukoyama A., Matsuno S., Okabe N.;

RT "Characterization of hexon and fiber genes of a novel strain of adenovirus involved in epidemic keratoconjunctivitis,";

RL J. Clin. Pathol. 57:95-97(2004).

DR EMBL; AB079724; BAC07480.1; -.

DR InterPro; IPR009266; Adeno_E3.

DR Pfam; PF06040; Adeno_E3; 1.

SQ SEQUENCE 131 AA; 14966 MW; 6EAE3911AADS1FBD CRC64;

Query Match 55.6%; Score 50; DB 2; Length 131;

Best Local Similarity 56.2%; Pred. No. 6.6;

Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GELEGLPIPNPLRTG 17
|||||:|

Db 77 GELEGLPIPNPLRTG 92
|||||:|

RESULT 2

Q8B8S5 ID Q8B8S5 PRELIMINARY; PRT; 131 AA.

AC Q8B8S5; DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 05-JUN-2004 (TREMBlrel. 27, Last annotation update)

DE E3 14.6 kDa protein (16.1K protein) (E3 CRI-alpha).

GN Name=E3A;

OS Human adenovirus type 11.

OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.

OX NCBI_TaxID=10541;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Adilp Slobitski;

RX MEDLINE=22614046; PubMed=12726735; DOI=10.1016/S0042-6822(02)00085-5;

RA Stone D., Furthmann A., Sandig V., Lieber A.;

RT "The complete nucleotide sequence, genome organization, and origin of human adenovirus type 11,";

RL Virology 309:152-165(2003).


```

RA  Bout A., Goudsmit J., Havenga M.;
RT  "Replication-deficient human adenovirus type 35 vectors for gene
RT  transfer and vaccination: efficient human cell infection and bypass of
RT  preexisting adenovirus immunity.";
RL  J. Virol. 77:8263-8271(2003).
DR  EMBL; AY128664; AAA75324.1; -.
DR  EMBL; AY128640; AANI7492.1; -.
DR  EMBL; AY271307; AAP92358.1; -.
DR  FIRM; JC4767; JC4767.
DR  InterPro; IPR009266; Adeno_E3.
DR  Pfam; PF06040; Adeno_E3; 1.
SQ  SEQUENCE 131 AA; 14974 MW; E4869881FC9F743 CRC64;

Query Match      53.3%; Score 48; DB 2; Length 131;
Best Local Similarity 56.2%; Pred. No. 14;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY  2 GELEGLPIPNPLRTG 17
    ||| ||| ||| :
DB  77 GELGLTPTNPWVZAG 92

RESULT 5
Q8P4Q6 PRELIMINARY; PRT; 154 AA.
AC  Q8P4Q6;
DT  01-OCT-2002 (TrEMBLrel. 22, Created)
DT  01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  Hypothetical protein XCC3651.
GN  OrderedLocusNames=XCC3651;
OS  Xanthomonas campestris (pv. campestris).
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC  Xanthomonadaceae; Xanthomonas.
OX  NCBI_TaxID=340;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=ATCC 33913 / NCPPB 528;
RX  MEDLINE=2022145; PubMed=12024217; DOI=10.1038/417459a;
RA  da Silva A.C.B., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA  Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA  Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA  Camarotte G., Cannavan F., Cardoso J., Chambergo F., Cipina L.P.,
RA  Ciccarelli R.M.B., Coutinho L.B., Cursino-Santos J.R., El-Dorri H.,
RA  Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA  Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA  Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA  Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA  Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA  Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA  Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA  Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA  Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA  Setubal J.C., Kitajima J.P.;
RT  "Comparison of the genomes of two Xanthomonas pathogens with differing
RT  host specificities.";
RL  Nature 417:459-463(2002).
DR  EMBL; AE012486; AAM42921.1; -.
DR  HSP; P14930; IL1D.
DR  InterPro; IPR002579; MsrB.
DR  InterPro; IPR011057; Msr4_like.
DR  Pfam; PF01641; SelR; 1.
DR  ProDom; PD004057; DUF25; 1.
DR  TIGRFAMs; TIGR00357; DUF25; 1.
KW  Complete proteome.
SQ  SEQUENCE 154 AA; 16794 MW; 44390679394D7D1F CRC64;

Query Match      53.3%; Score 48; DB 2; Length 154;
Best Local Similarity 61.5%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY  5 EGLPIPNPLRTG 17
    : ||| ||| |||

```

```

Db 136 DQQLPPLPQAG 148

RESULT 6
Q94HT5 PRELIMINARY; PRT; 147 AA.
AC  Q94HT5;
DT  01-DEC-2001 (TrEMBLrel. 19, Created)
DT  01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT  01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE  Putative retroelement.
GN  Name=OSJNB0036B06.22;
DN  Name=OSJNB0036B06.22;
OS  Oryza sativa (Rice).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC  Ehrhartoideae; Oryzaceae; Oryza.
OX  NCBI_TaxID=4530;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Nipponbare;
RA  Du H., Minx P., Abbott A., Doebber A., de la Bastide M., Spiegel L.,
RA  Nascimento L., Preston R., Kirchoff K., King L., Vil M.D., Baker J.,
RA  Zutavern T., Santos L., Bell M., Miller B., Kuit K., Rodriguez S.,
RA  Cunnius D.M., Baliya V., Shah R., Bahret A., O'Shaughnessy A.,
RA  Palmer L., Yang C., Dedhia N., McCombie W.R.;
RL  Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AC051624; AAK92564.1; -.
DR  Gramene; Q94HT5; -.
SQ  SEQUENCE 147 AA; 16177 MW; 7CF958DBE551325 CRC64;

Query Match      52.2%; Score 47; DB 2; Length 147;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY  1 KGELEGLPIPNPLR 15
    ||| ||| : |||
DB  35 KGLGLPVRSPTLR 49

RESULT 7
Q7XG14 PRELIMINARY; PRT; 147 AA.
AC  Q7XG14;
DT  01-OCT-2003 (TrEMBLrel. 25, Created)
DT  01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  Putative retroelement.
GN  ORFNames=OSJNB0036B06.22;
OS  Oryza sativa (japonica cultivar-group).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC  Ehrhartoideae; Oryzaceae; Oryza.
OX  NCBI_TaxID=39947;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  The Rice Chromosome 10 Sequencing Consortium;
RT  "In-depth view of structure, activity, and evolution of rice
RT  chromosome 10.";
RL  Science 300:1566-1569(2003).
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL  Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AE017073; AAP52859.1; -.
DR  Gramene; Q7XG14; -.
SQ  SEQUENCE 147 AA; 16177 MW; 7CF958DBE551325 CRC64;

Query Match      52.2%; Score 47; DB 2; Length 147;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY  1 KGELEGLPIPNPLR 15
    ||| ||| : |||

```

[illegible]

```

RL Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001089; AAB90990.1; -.
DR PIR; C69280; C69280.
DR TIGR; AF0243; -.
DR InterPro; IPR006638; Elp3/Miab/NiFb.
DR InterPro; IPR007197; Radical_SAM.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; Elp3; 1.
DR Complete proteome; Hypothetical protein.
KW SEQUENCE 529 AA; 60503 MW; 57E9FBC20D284BBF CRC64;

Query Match 52.2%; Score 47; DB 1; Length 529;
Best Local Similarity 72.7%; Pred. No. 88;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GELEGLPIPNP 12
Db 222 GEFEGVPRNP 232
||| ||| : |||
||| ||| : |||

RESULT 11
Q94HML PRELIMINARY; PRT; 579 AA.
AC Q94HML
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative retroelement.
GN Name=OSUNBa0089D15.27;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nipponbare;
RA Spiegel L., de la Bastide M., Nascimento L., Kirchhoff K., King L.,
RA Preston R., Vil M.D., Baker J., Bell M., Zutavern T., Santos L.,
RA Miller B., Kuit K., Rodriguez S., Cunnill D.M., Balija V., Shah R.,
RA Bahret A., Bai H., O'Shaughnessy A., Dedhia N., McCombie W.R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AC078944; AAK92616.1; -.
DR Gramene; Q94HML; -.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR005162; Retrotrans_gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
SQ SEQUENCE 579 AA; 63293 MW; 1832D98496BE25EF CRC64;

Query Match 52.2%; Score 47; DB 2; Length 579;
Best Local Similarity 60.0%; Pred. No. 97;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KGELEGLPIPNPLR 15
Db 35 KGLGLPVRSPTLR 49
||| ||| : |||
||| ||| : |||

RESULT 12
Q7XGI5 PRELIMINARY; PRT; 579 AA.
AC Q7XGI5
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

```

```

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative retroelement.
GN ORFNames=OSUNBa0089D15.27;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
RL chromosome 10.";
RL Science 300:1566-1569 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
EMBL; AE017061; AAP52277.1; -.
DR Gramene; Q7XGI5; -.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR005162; Retrotrans_gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
SQ SEQUENCE 579 AA; 63293 MW; 1832D98496BE25EF CRC64;

Query Match 52.2%; Score 47; DB 2; Length 579;
Best Local Similarity 60.0%; Pred. No. 97;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KGELEGLPIPNPLR 15
Db 35 KGLGLPVRSPTLR 49
||| ||| : |||
||| ||| : |||

RESULT 13
Q7S3G9 PRELIMINARY; PRT; 2039 AA.
AC Q7S3G9
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU06884.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehm B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Iankiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamvesselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofora S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000409; EAA29995.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR000379; Ser_estrs.
KW Hypothetical protein.

```

```
SQ SEQUENCE 2039 AA; 220637 MW; 62E6EAA1E02A38A3 CRC64;
Query Match 52.2%; Score 47; DB 2; Length 2039;
Best Local Similarity 64.3%; Pred. No. 3.7e+02;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 LEGLELPINPLRTG 17
Db 1178 LDGLEIPTPALEGT 1191

RESULT 14
Q6MND1 PRELIMINARY; PRT; 202 AA.
ID Q6MND1
AC Q6MND1
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=Bdi1327;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goessmann A., Meyer F.,
RA Sackett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective.";
RL Science 303:689-692(2004)
DR EMBL; BX842649; CAE79221.1; -.
KW Complete proteome.
SQ SEQUENCE 202 AA; 22573 MW; 35DBF670C9A4972E CRC64;

Query Match 51.1%; Score 46; DB 2; Length 202;
Best Local Similarity 58.3%; Pred. No. 45;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GELEGLPVPNPL 13
Db 147 GFEGGLPVGVPNI 158

RESULT 15
Q8UGI3 PRELIMINARY; PRT; 255 AA.
ID Q8UGI3
AC Q8UGI3; Q7D026;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein Atu1054 (AGR_C1946p).
GN OrderedLocusNames=AGR_C1946, Atu1054;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dupont;
RX MEDLINE=2168550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo Y., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
```

```
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Cereon;
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorllo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger W., Doughty D., Scott C., Leppas C., Markelz B.,
RA Planagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE009069; AAL42067.1; -.
DR EMBL; AE008035; AAK86862.1; -.
DR PIR; AE2706; AE2706.
DR PIR; E97488; E97488.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 255 AA; 27291 MW; 5B329E22B9C43683 CRC64;

Query Match 51.1%; Score 46; DB 2; Length 255;
Best Local Similarity 42.9%; Pred. No. 58;
Matches 6; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 KGELEGLPVPNPL 14
Db 68 EGDSDGIPLPDPLV 81

Search completed: May 27, 2005, 13:47:30
Job time : 6.73348 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:34:21 ; Search time 1.3049 Seconds
(without alignments)
1253.491 Million cell updates/sec

Title: US-10-622-893a-4
Perfect score: 90
Sequence: 1 KGEGLPIPNPLRTG 17
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	55.6	529	2 G64131	sporulation protei
2	48	53.3	131	2 JC4767	E3 gene encoding h
3	47	52.2	358	2 A54265	CCAAT/enhancer-bin
4	47	52.2	529	2 C69280	conserved hypotet
5	46	51.1	255	2 A82706	conserved hypotet
6	46	51.1	255	2 E97488	hypothetical prote
7	45	50.0	378	2 AF3320	flavohemoprotein [
8	45	50.0	560	2 S51600	phosphorylase kina
9	45	50.0	757	2 A39283	gamma-glutamyl car
10	45	50.0	758	2 A47439	gamma-glutamyl car
11	45	50.0	1371	2 S77521	sensory transducti
12	44	48.9	103	1 SXADC2	hexon-associated p
13	44	48.9	145	2 G70384	hypothetical prote
14	44	48.9	266	2 S44954	lmbG protein - Str
15	43.5	48.3	360	2 H87699	conserved hypotet
16	43.5	48.3	768	2 T17422	disease resistance
17	43	47.8	175	2 H72538	hypothetical prote
18	43	47.8	222	1 MNZSP	nonstructural prot
19	43	47.8	358	2 H75264	hypothetical prote
20	43	47.8	392	1 RRNZSP	polymerase-associ
21	43	47.8	820	2 T45685	hypothetical prote
22	43	47.8	855	2 B75191	probable ATP-depen
23	42.5	47.2	437	2 C39135	hypothetical prote
24	42	46.7	253	2 C81670	undecaprenyl pyrop
25	42	46.7	253	2 E71512	hypothetical prote
26	42	46.7	319	2 AG0700	probable pathogeni
27	42	46.7	322	2 E82505	conserved hypotet
28	42	46.7	473	2 S70357	forkhead transcrip
29	42	46.7	598	2 F83977	hypothetical prote

30	42	46.7	612	2 T13616	hypothetical prote
31	42	46.7	778	2 A23308	transcription init
32	42	46.7	1071	1 FAXIVA	H+-exporting ATPas
33	42	46.7	1998	2 T08822	nonstructural poly
34	41	45.6	264	2 S77408	phosphoribosylamin
35	41	45.6	269	2 AH1970	type 4 prepilin pe
36	41	45.6	302	1 H64127	ribosomal protein
37	41	45.6	335	2 E75183	ATP-binding protei
38	41	45.6	342	2 A10286	pap operon transcr
39	41	45.6	421	2 A12824	NADH dehydrogenase
40	41	45.6	438	2 H97602	probable NADH dehy
41	41	45.6	463	2 F90422	4-hydroxybutyryl-C
42	41	45.6	514	2 T43304	Rad17 protein homo
43	41	45.6	579	2 C71916	aspartate-tRNA lig
44	41	45.6	606	2 T27035	hypothetical prote
45	41	45.6	772	2 S22660	excision repair pr

ALIGNMENTS

RESULT 1

G64131
sporulation protein spoIIIE - Haemophilus influenzae (strain Rd KW20)
C/Species: Haemophilus influenzae
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C/Accession: G64131
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: G64131
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-529 <TIGR>
A;Cross-references: UNIPROT:P45264; GB:U32833; GB:L42023; NID:g1574432; PIDN:AAC23240.1;
C;Superfamily: Bacillus subtilis DNA translocase spoIIIE

Query Match 55.6%; Score 50; DB 2; Length 529;
Best Local Similarity 53.3%; Pred. No. 4.3;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 ELEGLEPIPNPLRTG 17
| | | | | | | | | |
Db 286 EAMGMPVPNPIWRIG 300

RESULT 2

JC4767
E3 gene encoding hypothetical 15.0k protein - human adenovirus 35
C/Species: Mastadenovirus h35 (human adenovirus 35)
C/Date: 10-May-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C/Accession: JC4767
R;Basler, C.F.; Droguett, G.; Horwitz, M.S. Gene 170, 249-254, 1996
A;Title: Sequence of the immunoregulatory early region 3 and flanking sequences of adeno
A;Reference number: JC4765; MUID:96235144; PMID:8666254
A;Accession: JC4767
A;Molecule type: DNA
A;Residues: 1-131 <BAS>
A;Cross-references: UNIPROT:Q67729; GB:U32664; NID:g984529; PIDN:AAA75324.1; PID:g984532
C;Genetics:
A;Gene: E3
C;Superfamily: adenovirus early E3 16K glycoprotein

Query Match 53.3%; Score 48; DB 2; Length 131;
Best Local Similarity 56.2%; Pred. No. 1.8;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GELEGLEPIPNPLRTG 17

Db 77 GELHGLPTENPWVEAG 92

RESULT 3

AS4265
 CCAAT/enhancer-binding protein - rat
 N;Alternate names: C/EBP
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 06-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
 C;Accession: S06890; A54265
 R;Landschulz, W.H.; Johnson, P.F.; Adashi, E.Y.; Graves, B.J.; McKnight, S.L.
 Genes Dev. 2, 786-800, 1988
 A;Title: Isolation of a recombinant copy of the gene encoding C/EBP.
 A;Reference number: S06890; MUID:89092001; PMID:2850264
 A;Accession: S06890
 A;Molecule type: DNA
 A;Residues: 1-29, RWLSPGRPRAAPSTCRPGNAGR', 55-358 <LA>
 A;Cross-references: UNIPROT:P05554; EMBL:X12752
 A;Note: the authors translated the codon AGC for residue 21 as Asp
 A;Note: part of this sequence was confirmed by protein sequencing
 A;Note: this sequence has been corrected
 R;Lincoln, A.J.; Williams, S.C.; Johnson, P.F.
 Genes Dev. 8, 1131-1132, 1994
 A;Title: A revised sequence of the rat C/EBP gene.
 A;Reference number: A54265; MUID:95011606; PMID:7926792
 A;Accession: A54265
 A;Molecule type: DNA
 A;Residues: 1-60 <LIN>
 A;Cross-references: GB:X12752
 C;Superfamily: CCAAT/enhancer-binding protein alpha
 C;Keywords: transcription factor

Query Match 52.2%; Score 47; DB 2; Length 358;
 Best Local Similarity 62.5%; Pred. No. 8.4;
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GELEGLPIPNPLRTG 17
 |||||
 Db 247 GSLKGLAGHPDLRTG 262
 |||||

RESULT 4

C69280
 conserved hypothetical protein AF0243 - Archaeoglobus fulgidus
 C;Species: Archaeoglobus fulgidus
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C;Accession: C69280
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A;Reference number: A69250; MUID:98049343; PMID:9389475
 A;Accession: C69280
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-529 <KLE>
 A;Cross-references: UNIPROT:Q29996; GB:AE000782; MID:g2689412; PIDN:AAB9099
 C;Superfamily: Pyrococcus abyssi hypothetical protein PAB0059

Query Match 52.2%; Score 47; DB 2; Length 529;
 Best Local Similarity 72.7%; Pred. No. 13;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GELEGLPIPNP 12
 |||||
 Db 222 GEFEGVPRNP 232
 |||||

RESULT 5

AE2706
 conserved hypothetical protein Atul054 [imported] - Agrobacterium tumefaciens (strain C5
 C;Species: Agrobacterium tumefaciens
 C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C;Accession: AE2706
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCella
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I.
 ster, E.W.
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A;Reference number: AB2577; MUID:21608550; PMID:11743193
 A;Accession: AE2706
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-255 <KUR>
 A;Cross-references: UNIPROT:Q8UGI3; GB:AE008688; PIDN:AAL42067.1; PID:g17739446; GSPDB:G
 A;Experimental source: strain C58 (Dupont)
 C;Genetics:
 A;Gene: Atul054
 A;Map position: circular chromosome

Query Match 51.1%; Score 46; DB 2; Length 255;
 Best Local Similarity 42.9%; Pred. No. 8.2;
 Matches 6; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KGELEGLPIPNPLL 14
 :||:|:|:|:|:
 Db 68 EGSDGIPLPDPLV 81
 :||:|:|:|:|:

RESULT 6

E97488
 hypothetical protein AGR_C_1946 [imported] - Agrobacterium tumefaciens (strain C58, Cere
 C;Species: Agrobacterium tumefaciens
 C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C;Accession: E97488
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A;Reference number: A97359; MUID:21608551; PMID:11743194
 A;Accession: E97488
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-255 <KUR>
 A;Cross-references: UNIPROT:Q8UGI3; GB:AE007869; PIDN:AAK86862.1; PID:g15156078; GSPDB:G
 C;Genetics:
 A;Gene: AGR_C_1946
 A;Map position: circular chromosome

Query Match 51.1%; Score 46; DB 2; Length 255;
 Best Local Similarity 42.9%; Pred. No. 8.2;
 Matches 6; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KGELEGLPIPNPLL 14
 :||:|:|:|:|:
 Db 68 EGSDGIPLPDPLV 81
 :||:|:|:|:|:

RESULT 7

AF3320
 flavohemoprotein [imported] - Brucella melitensis (strain 16M)
 C;Species: Brucella melitensis
 C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 16-Aug-2004
 C;Accession: AF3320
 R;DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, I
 ; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesc
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A;Reference number: AD3252; PMID:11756688
 A;Accession: AF3320

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-378 <KUR>
A:Cross-references: UNIPROT:Q8YI97; GB:AE008917; PIDN:AAL51729.1; PID:g17982465; GSPDB:C
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0548
A:Map position: 1
C:Superfamily: Phthalate dioxygenase reductase; cytochrome-b5 reductase homology; ferred
  Query Match 50.0%; Score 45; DB 2; Length 378;
  Best Local Similarity 46.7%; Pred. No. 19;
  Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 ELEGLEPIPNPLRTG 17
   |:|:|:|:|
Db 266 EISAVPVSPPLPETG 280

RESULT 8
S51600
Phosphorylase kinase (EC 2.7.1.38) gamma chain - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 16-Aug-2004
C:Accession: S51600
R:Bahri, S.M.; Chia, W.
Mol. Gen. Genet. 245, 588-597, 1994
A:Title: DPHK-gamma, a putative Drosophila kinase with homology to vertebrate phosphoryl
  utants.
A:Reference number: S51600; MUID:95107257; PMID:7808409
A:Accession: S51600
A:Molecule type: DNA
A:Residues: 1-560 <BAH>
A:Cross-references: UNIPROT:Q9YVV7; EMBL:U13014; NID:9531470; PIDN:AAA64560.1; PID:95314
A:Note: the authors translated the codon CAT for residue 528 as Ala and CAC for residue
C:Genetics:
A:Gene: FlyBase:PhK-gamma
A:Cross-references: FlyBase:FBgn0011754
A:Map position: 10
C:Superfamily: protein kinase homology
C:Keywords: phosphotransferase
P:21-291/Domain: protein kinase homology <KIN>

  Query Match 50.0%; Score 45; DB 2; Length 560;
  Best Local Similarity 72.7%; Pred. No. 30;
  Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 7 LPINPLLRGT 17
   |:|:|:|:|
Db 439 LQLNPPLLTG 449

RESULT 9
A39283
gamma-glutamyl carboxylase (EC 4.1.1.-) - human
C:Species: Homo sapiens (man)
C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 06-Dec-1996
C:Accession: A39283
R:Wu, S.M.; Cheung, W.F.; Frazier, D.; Stafford, D.W.
Science 254, 1634-1636, 1991
A:Title: Cloning and expression of the cDNA for human gamma-glutamyl carboxylase.
A:Reference number: A39283; MUID:92086858; PMID:1749935
A:Accession: A39283
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-757 <WUA>
A:Cross-references: GB:M81592
C:Keywords: carbon-carbon lyase; carboxy-lyase

  Query Match 50.0%; Score 45; DB 2; Length 757;
  Best Local Similarity 50.0%; Pred. No. 42;
  Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 7 LPINPLLRGT 17
   |:|:|:|:|
Db 439 LQLNPPLLTG 449

RESULT 9
S77521
sensory transduction histidine kinase sll1124 - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein sll1124
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S77521
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
  O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
  DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
  s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S77521
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1371 <KAN>
A:Cross-references: UNIPROT:P73337; EMBL:D90905; GB:AB001339; NID:g1652360; PIDN:BA01736
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

  Query Match 50.0%; Score 45; DB 2; Length 1371;
  Best Local Similarity 50.0%; Pred. No. 82;

```

```

Qy 1 KGELEGLPIPNPLRT 16
   |:|:|:|:|
Db 643 EGEVKGPEPTPLVQT 658

RESULT 10
A47439
gamma-glutamyl carboxylase (EC 4.1.1.-) - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: A47439; A38939; B39283; C39283
R:Rehemulla, A.; Roth, D.A.; Wasley, L.C.; Kuliopulos, A.; Walsh, C.T.; Furie, B.; Furi
  Proc. Natl. Acad. Sci. U.S.A. 90, 4611-4615, 1993
A:Title: In vitro and in vivo functional characterization of bovine vitamin K-dependent
  A:Reference number: A47439; MUID:93281629; PMID:8506307
A:Accession: A47439
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-758 <REH>
A:Cross-references: UNIPROT:Q07175; GB:L09726; NID:g289398; PIDN:AAA30425.1; PID:g289399
R:Wu, S.M.; Cheung, W.F.; Frazier, D.; Stafford, D.W.
submitted to Genbank, January 1992
A:Reference number: A38939
A:Accession: A38939
A:Molecule type: mRNA
A:Residues: 96-758 <WUA>
A:Cross-references: GB:M81593; NID:g162744; PIDN:AAA30410.1; PID:g162745
R:Wu, S.M.; Cheung, W.F.; Frazier, D.; Stafford, D.W.
Science 254, 1634-1636, 1991
A:Title: Cloning and expression of the cDNA for human gamma-glutamyl carboxylase.
A:Reference number: A39283; MUID:92086858; PMID:1749935
A:Accession: B39283
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 96-693,695-758 <WU2>
A:Cross-references: GB:M81593
A:Accession: C39283
A:Molecule type: protein
A:Residues: 37-50 <WU3>
C:Keywords: carbon-carbon lyase; carboxy-lyase

  Query Match 50.0%; Score 45; DB 2; Length 758;
  Best Local Similarity 50.0%; Pred. No. 42;
  Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KGELEGLPIPNPLRT 16
   |:|:|:|:|
Db 643 EGEVKGPEPTPLVQT 658

RESULT 11
S77521
sensory transduction histidine kinase sll1124 - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein sll1124
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S77521
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
  O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
  DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
  s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S77521
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1371 <KAN>
A:Cross-references: UNIPROT:P73337; EMBL:D90905; GB:AB001339; NID:g1652360; PIDN:BA01736
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

  Query Match 50.0%; Score 45; DB 2; Length 1371;
  Best Local Similarity 50.0%; Pred. No. 82;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:47:36 ; Search time 136.87 Seconds
(without alignments)
1189.648 Million cell updates/sec

Title: US-10-622-893A-5

Perfect score: 2530

Sequence: 1 MGSGDDDDALAVTKSSSL.....LEGIPNPILRTGHHHHH 472

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1462099 seqs, 344972447 residues

Total number of hits satisfying chosen parameters: 1462099

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_PUBCOMB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2530	100.0	472	17	US-10-622-893A-5
2	2329	92.1	437	17	US-10-622-893A-3
3	1153.5	45.6	440	17	US-10-470-678-9
4	1152.5	45.6	440	17	US-10-470-678-10
5	1151.5	45.5	440	17	US-10-470-678-8
6	627	24.8	437	14	US-10-232-655-3
7	600	23.7	437	14	US-10-232-655-1
8	210.5	8.3	139	16	US-10-767-701-53616
9	184	7.3	384	14	US-10-156-761-14478
10	168	6.6	433	14	US-10-287-274-412
11	168	6.6	433	15	US-10-282-122A-42701
12	166	6.6	387	17	US-10-829-427-1
13	165	6.5	385	15	US-10-282-122A-68225

14	159.5	6.3	381	14	US-10-156-761-9328	Sequence 9328, Ap
15	140	5.5	374	15	US-10-282-122A-55515	Sequence 55515, A
16	137.5	5.4	420	15	US-10-425-114-54552	Sequence 54552, A
17	136.5	5.4	372	15	US-10-282-122A-43213	Sequence 43213, A
18	136.5	5.4	383	9	US-09-883-096-5	Sequence 5, Appli
19	135.5	5.4	378	10	US-09-975-719-367	Sequence 367, App
20	127.5	5.0	433	9	US-09-815-242-13745	Sequence 13745, A
21	127	5.0	372	15	US-10-282-122A-75876	Sequence 75876, A
22	127	5.0	414	16	US-10-437-963-177242	Sequence 177242, A
23	126	5.0	370	15	US-10-282-122A-72847	Sequence 72847, A
24	125.5	5.0	570	17	US-10-800-350-386	Sequence 386, App
25	124	4.9	429	15	US-10-425-114-71622	Sequence 71622, A
26	122	4.8	406	17	US-10-485-710-80	Sequence 80, Appl
27	121.5	4.8	669	9	US-09-213-888-25	Sequence 25, Appl
28	121.5	4.8	669	9	US-09-328-877A-25	Sequence 25, Appl
29	121.5	4.8	669	16	US-10-653-497-25	Sequence 25, Appl
30	121.5	4.8	669	16	US-10-653-517-25	Sequence 25, Appl
31	121.5	4.8	669	16	US-10-653-496A-25	Sequence 25, Appl
32	121.5	4.8	669	17	US-10-653-676A-25	Sequence 25, Appl
33	120.5	4.8	405	14	US-10-156-761-9148	Sequence 9148, Ap
34	120	4.7	818	14	US-10-156-761-14479	Sequence 14479, A
35	119.5	4.7	391	15	US-10-282-122A-69762	Sequence 69762, A
36	115.5	4.6	568	10	US-09-953-348-135	Sequence 135, App
37	115.5	4.6	568	14	US-10-267-255-135	Sequence 135, App
38	115	4.5	563	15	US-10-415-302-8	Sequence 8, Appli
39	115	4.5	563	15	US-10-415-302-18	Sequence 18, Appl
40	113.5	4.5	395	9	US-09-873-880-30	Sequence 30, Appl
41	113.5	4.5	395	15	US-10-424-599-268694	Sequence 268694, A
42	113.5	4.5	542	16	US-10-437-963-117091	Sequence 117091, A
43	113.5	4.5	632	15	US-10-282-122A-68951	Sequence 68951, A
44	113.5	4.5	852	15	US-10-282-122A-50546	Sequence 50546, A
45	112.5	4.4	317	15	US-10-282-122A-61710	Sequence 61710, A

ALIGNMENTS

RESULT 1
US-10-622-893A-5
; Sequence 5, Application US/10622893A
; Publication No. US20050014935A1
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong-Sheng
; APPLICANT: Datta, Abhijit
; APPLICANT: Wang, Yuping
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: DETERMINATION OF GLYCATED PROTEINS
; FILE REFERENCE: 46692001300
; CURRENT APPLICATION NUMBER: US/10/622,893A
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric protein
US-10-622-893A-5

Query Match	100.0%	Score	2530	DB	17	Length	472
Best Local Similarity	100.0%	Pred. No.	7.6e-237				
Matches	472	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MGSGDDDDALAVTKSSSL	1	MGSGDDDDALAVTKSSSL	1	MGSGDDDDALAVTKSSSL	1
Db	1	MGSGDDDDALAVTKSSSL	1	MGSGDDDDALAVTKSSSL	1	MGSGDDDDALAVTKSSSL	1
Qy	61	DVNRVSSGQYNNKDI	61	DVNRVSSGQYNNKDI	61	DVNRVSSGQYNNKDI	61
Db	61	DVNRVSSGQYNNKDI	61	DVNRVSSGQYNNKDI	61	DVNRVSSGQYNNKDI	61
Qy	121	GVRVRPEDNLVELT	121	GVRVRPEDNLVELT	121	GVRVRPEDNLVELT	121
Db	121	GVRVRPEDNLVELT	121	GVRVRPEDNLVELT	121	GVRVRPEDNLVELT	121

Db 121 GVRVREGEDNVLVELTRPEQFRKLAPEGVLQDFPGWKGYFARSGAGWAHARNALVAAAR 180
Qy 181 EAQRMGVKFTGTPQGRVVTLIFENNNDVGAVTGDKIWAERTFLCAGASAGQFLDFKN 240
Db 181 EAQRMGVKFTGTPQGRVVTLIFENNNDVGAVTGDKIWAERTFLCAGASAGQFLDFKN 240
Qy 241 QLRPTAWTLVHIALKPEERALKYKNIPVIFNIERGFPEDEERGEIKI CDEHPGYTNMVQ 300
Db 241 QLRPTAWTLVHIALKPEERALKYKNIPVIFNIERGFPEDEERGEIKI CDEHPGYTNMVQ 300
Qy 301 SADGTWMSIPFEKTQIPKEAETRVALLKETMPOLADRPFSFARICWCADTANREFLIDR 360
Db 301 SADGTWMSIPFEKTQIPKEAETRVALLKETMPOLADRPFSFARICWCADTANREFLIDR 360
Qy 361 HPQYHSLVLCGASGRGFKPLPSIGNLI VDAMEGKVPQKIHELIIKNPDI AARNNRWDTL 420
Db 361 HPQYHSLVLCGASGRGFKPLPSIGNLI VDAMEGKVPQKIHELIIKNPDI AARNNRWDTL 420
Qy 421 GRFGGPNRVMDFDHVKWNTNVQYRDISKLKLEGLPIPNPLLRTHGHHHH 472
Db 421 GRFGGPNRVMDFDHVKWNTNVQYRDISKLKLEGLPIPNPLLRTHGHHHH 472

RESULT 2
US-10-622-893A-3
; Sequence 3, Application US/10622893A
; Publication No. US20050014935A1
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong-Sheng
; APPLICANT: Datta, Abhijit
; APPLICANT: Wang, Yuping
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: DETERMINATION OF GLYCATED PROTEINS
; FILE REFERENCE: 466992001300
; CURRENT APPLICATION NUMBER: US/10/622, 893A
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 40%-100% identity to the amadoriase
US-10-622-893A-3

Query Match 92.1%; Score 2329; DB 17; Length 437;
Best Local Similarity 100.0%; Pred. No. 2.4e-217;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 AVTKSSLLIVGAGTGTSTALHARRGYTNVTVDLPYPVPSAISAGNDVNVKISSGOYS 72
Db 1 AVTKSSLLIVGAGTGTSTALHARRGYTNVTVDLPYPVPSAISAGNDVNVKISSGOYS 60
Qy 73 NNKDEIENVEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQGLDRLGVVRPGE DPNL 132
Db 61 NNKDEIENVEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQGLDRLGVVRPGE DPNL 120
Qy 133 VELTRPEQFRKLAPEGVLQDFPGWKGYFARSGAGWAHARNALVAAAREAQRGVKPV 192
Db 121 VELTRPEQFRKLAPEGVLQDFPGWKGYFARSGAGWAHARNALVAAAREAQRGVKPV 180
Qy 193 TPQGRVVTLIFENNNDVGATGDKIWAERTFLCAGASAGQFLDFKNLRPTAWTLVHI 252
Db 181 TPQGRVVTLIFENNNDVGATGDKIWAERTFLCAGASAGQFLDFKNLRPTAWTLVHI 240
Qy 253 ALKPEERALKYKNIPVIFNIERGFPEDEERGEIKI CDEHPGYTNMVQSDGTWMSIPFE 312
Db 241 ALKPEERALKYKNIPVIFNIERGFPEDEERGEIKI CDEHPGYTNMVQSDGTWMSIPFE 300
Qy 313 KTIQIPKEAETRVALLKETMPOLADRPFSFARICWCADTANREFLIDRHPQYHSLVLCG 372

Db 301 KTIQIPKEAETRVALLKETMPOLADRPFSFARICWCADTANREFLIDRHPQYHSLVLCG 360
Qy 373 ASGRGFKYLPISIGNLI VDAMEGKVPQKIHELIIKNPDI AARNNRWDTLGRFGGPNRVMD 432
Db 361 ASGRGFKYLPISIGNLI VDAMEGKVPQKIHELIIKNPDI AARNNRWDTLGRFGGPNRVMD 420
Qy 433 HDVKWNTNVQYRDISKL 449
Db 421 HDVKWNTNVQYRDISKL 437

RESULT 3
US-10-470-678-9
; Sequence 9, Application US/10470678
; Publication No. US20050101711A1
; GENERAL INFORMATION:
; APPLICANT: KOUZUMA, Takuji et al.
; TITLE OF INVENTION: COMPOSITION FOR ASSAYING GLYCATED PROTEINS
; FILE REFERENCE: 1516-0121P
; CURRENT APPLICATION NUMBER: US/10/470, 678
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: JP 2001/ 22953
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: JP 2001/ 39796
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: JP 2001/ 24002
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic protein sequence of SEQ ID NO: 3
; NAME/KEY: MUTAGEN
; LOCATION: (372)..(372)
US-10-470-678-9

Query Match 45.6%; Score 1153.5; DB 17; Length 440;
Best Local Similarity 50.9%; Pred. No. 5.6e-103;
Matches 223; Conservative 75; Mismatches 133; Indels 7; Gaps 5;

Qy 11 ALAVTKSSLLIVGAGTGTSTALHARRGYTNVTVDLPYPVPSAISAGNDVNVKIS--S 68
Db 1 ASLTITKQSQILLVGGTGWGCTALHARRGYTNVTVDVNRIPSPISAGHDVNVKLAGRLS 60
Qy 69 GOYSNNKDEIENVEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQGLDRLGVVRPGE 128
Db 61 TADSKGDDSDSIWKALSAAAAQGLHDPVFPQPCHTGTSVAVGSTPKSIKQL-VEDEIGD 119
Qy 129 D-PNLVELTRPEQFRKLAPEGVLQDFPGWKGYFARSGAGWAHARNALVAAAREAQRMGV 187
Db 120 DIDQYTPLTAAEDFRKTMPEGILTNFPGWKGYKPTGSGVHARKAMKRAAFESRELGV 179
Qy 188 KFTVGTPOGRVVTLIFENNNDVGAVTGDKIWAERTFLCAGASAGQFLDFKNQLRPTAW 247
Db 180 KFTGSPGKVESLIEDGVDVGAKTADGKEHADRTILSAGASAEFFLDQFNQIOPTAW 239
Qy 248 TLVHIALKPEERALKYKNIPVIFNIERGFPEDEERGEIKI CDEHPGYTNMVQSDGTM- 306
Db 240 TLGHIQMTPEETKLYKNLPFLFNINQGFMEPPEDLHQLKMCDEHPGCMNVKPGSKYP 299
Qy 307 MSTPFEKTQIPKEAETRVALLKETMPOLADRPFSFARICWCADTANREFLIDRHPQYHS 366
Db 300 QSIPTFAKHQVPTAEARMKQFLKDIIMPQLADRLPVHARICWCADTQORMFLIITHPHPS 359
Qy 367 LVLCGASGRGFKYLPISIGNLI VDAMEGKVPQKIHELIIKNPDI AARNNRWDTLGRFGG 426
Db 360 LVIASGDCGTGYMHITISIGKIFISDCMEGTLEERFAKFWRRPEKFTFEWKGKDLDRFGAD 419
Qy 427 NRVMDF--HDVKWNTNVQ 442


```
; CURRENT APPLICATION NUMBER: US/10/232,655
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: JP 2001-266665
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: JP 2001-378151
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: JP 2002-228727
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Eupenicillium terrenum
US-10-232-655-3

Query Match      24.8%; Score 627; DB 14; Length 437;
Best Local Similarity 34.1%; Pred. No. 9.2e-52;
Matches 153; Conservative 82; Mismatches 174; Indels 40; Gaps 14;

QY 12 LAVTKSSLLIV--GAGTGTSTALHLARRGYT--NVTVLDPYVPVSATISAGNDVNKVIS 67
Db 1 MAHSRASTKVVVGGGTGSSSTALHLIRSGYTPSNITVLDVYKTSLSQSAHDLNKNIM-- 59
QY 68 SGQYSNNKDEIEVNEILAEAEAFNGKNDPLFKPYHYDHTGLLMSACSQEGLDRLGVRVP- 126
Db 60 -GIRLRNGPDLQ---LSLESLDMMQNDLDFKPFHQVGMIDCSSKKEGLENLRRKYQTL 114
QY 127 -----GEDPNLVELTRPEQFKLAPGVLQGDPPGKGYFARGAGWAHARNALVAAARE 181
Db 115 LDAGIGLEKTNVWLESEDEILAKAP-NFTREQVKGWKLGFCTDG-GWLAATAKAINAIGIF 172
QY 182 AQRMGVKFVGTGPGQGVVTLIP--ENNVDKGVATGDKTWRABRTFLCAGASAGQFLDFK 239
Db 173 LQDKGVKFGFG-GAGTFQOPLFAADGKTGIGLETTDGTGKFADKVVLAAGAWSPTLVDL 231
QY 240 NQLRPTAWTLVHTALKPEERALKNIPIVFNIERGFFFEDEERGEIKICDEHPGYTNM- 298
Db 232 DQCVSKAWFAHQLTPKEADAYKNVPVYDGEVGFEEFN-EYGVIKVCDEFPFGSRFK 290
QY 299 -----VOSADGTMMISPEKTOIPKE-----AETRVRALLKETMPQADRPFSFARICWCA 349
Db 291 LHOPYGAASPKNISVPSRSHAKHPTDTPDASEVTIRKAIARFLPEFKDELFRNTMCWCT 350
QY 350 DTANREFLIDRHPQYHSLVLCGASGRGFKYLSIGNLIVDAMEGKVPQKIHILIKWNP 409
Db 351 DTADANLLICEHPKKNFILATGDSHGSFKILPNVGHVVELIEGRUPEEMAYQWRWRPGDALKSRRAAP 410
QY 410 IAAANRNRDTLGRFGGPNRVMDFHVKW 438
Db 411 GDALRS-----RRGAPAK--DLAEMPGW 431

RESULT 7
US-10-232-655-1
; Sequence 1, Application US/10232655
; Publication No. US20030157593A1
; GENERAL INFORMATION:
; APPLICANT: KUROSAWA, KEIKO
; APPLICANT: HIROKAWA, KOZO
; APPLICANT: KAJIYAMA, NAOKI
; TITLE OF INVENTION: No. US20030157593A1e1 fructosyl peptide oxidase
; FILE REFERENCE: 227590U0
; CURRENT APPLICATION NUMBER: US/10/232,655
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: JP 2001-266665
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: JP 2001-378151
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: JP 2002-228727
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Coniochaeta sp.
US-10-232-655-1

Query Match      23.7%; Score 600; DB 14; Length 437;
Best Local Similarity 32.7%; Pred. No. 3.9e-49;
Matches 141; Conservative 86; Mismatches 162; Indels 42; Gaps 13;

QY 20 LLIVGAGTGTSTALHLARRGY--TNVTVLDPYVPVSATISAGNDVNKVISGQYSNNKDE 77
Db 11 IVVGGGTGSSSTALHLVRSGYAPANITVLDTFEIPSAQSAGHDLNKNIM--GIRLRNKVD 68
QY 78 IEVNEILAEAEAFNGKNDPLFKPYHYDHTGLLMSACSQEGLDRLGVRVRPGEDPNL-VELT 136
Db 69 LQ-----MSLEARQMWKEDLFPQFFHNTGRMDCHEHTPEGIEDLKKQYQALHDAGAGLEKT 124
QY 137 R---PEQFKLAPGVLQGD--PPGKGYFARGAGWAHARNALVAAAREAQRMGVKEFTVG 192
Db 125 HAWLDNDEILSKMPLLRDQIQGWKAIWSQDG-GWLAATAKAINAIGQFLKERGVKFGFG 183
QY 193 TPQGRVVVTLIPENNVDK--GAVTGDGKIWRABRTFLCAGASAGQFLDFKNQLRPTATLV 250
Db 184 -GAGSKFQPLFDDGTTCTGIVETADGTXYADKVVLAAGAWSPTLVLDLDDQCCSKAWVYA 242
QY 251 HIALKPEERALKNIPIVFNIERGFFFEDEERGEIKICDEHPGYTNMVQ-----SADGT 305
Db 243 HQLTPEEAAAYKGVVNVNGSFGFFEPN-EFGVLKVCDEFPFGSRFKEHQPYGAPSPK 301
QY 306 MMSIPEKTOIPKE-----AETRVRALLKETMPQADRPFSFARICWCA DTANREFLIDR 360
Db 302 HISVPSRSHAKHPTDTPDASEVSIKKAIAITFLPRFQDKELFNRLALCWCTDTADAALIMCE 361
QY 361 HPQYHSLVLCGASGRGFKYLSIGNLIVDAMEGKVPQKIHILIKWNP----- 408
Db 362 HPKWKNFILATGDSHGSFKILPNVGHVVELIEGRUPEEMAYQWRWRPGDALKSRRAAP 421
QY 409 --DIAANRNR 417
Db 422 PKDLADMPGWK 432

RESULT 8
US-10-767-701-53616
; Sequence 53616, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 53616
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 13394327.pap
US-10-767-701-53616

Query Match      8.3%; Score 210.5; DB 16; Length 139;
Best Local Similarity 33.6%; Pred. No. 5.3e-12;
Matches 44; Conservative 21; Mismatches 39; Indels 27; Gaps 3;

QY 315 QIPKEAETRVRALLKETMPQADRPFSFARICWCA DTANREFLIDRHPQYHSLVLCGAS 374
Db 7 EIPREGQDVCREFLRCIPEFGNRPWTHRCWYTTDPTGWDWIDYHPKYKGLFVATGGS 66
```


APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA 034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 42701
LENGTH: 433
TYPE: PRT
ORGANISM: Escherichia coli
US-10-282-122A-42701

Query Match 6.6%; Score 168; DB 15; Length 433;
Best Local Similarity 23.0%; Pred. No. 4.1e-07;
Matches 120; Conservative 59; Mismatches 176; Indels 166; Gaps 28;
QY 1 MGSGDDDLALAVTKSSLLIVAGTGTCTSTALHARCYNTVTVLPVPVPSAISAGN 60
DB 7 LGHDMEDD-----CDIIITGAGTAGTACALRCARAG-LSVLLERAEIPG----- 50
QY 61 DVNKVSSGGYSNNKDBIEVNEILAEAFNGWKNDFLPKPYHDTGLLMSACSOEGLDRL 120
DB 51 --SKNLSGGRL-----YTHALAE-----LLPQFLHTAPLERRITHESLSLL 89
QY 121 ---GVRVRFGEDPNLVELTRPEQFRKLAPEG-----VLQGDFFGWKGYFARSGAGNAHARN 173
DB 90 TPDGVTT-----FSLQPGGESWSVLARFDFW----- 117
QY 174 ALVAAREQRMGKVFVTGPQGRVVTLLIFENNVDKAVTGDGKIWRARTFLCAGASAG 233
DB 118 -LVA---EAEKEGVECI---PGATVDALYEENGRCVIGCD--DILRARYVVLAEAG--- 166
QY 234 QFLDFKNLRPTAWTLVHIALKPEERALYKNIPIVFN-----TERGFFPEPDERGEI-- 286
DB 167 -----NSVLAERHGLV---TRPAGEAMALGIKEVLSLETSAEERPHLENNEGAALLFS 217
QY 287 -KICDBHPG---YTNVVOAGD-----TWMSIPFEKTOIPKEATRVALLKET-- 331
DB 218 GRICDDLPGGAFLYTQOQLSLGLVCPVLSLQSRVPASSELLTRFRAHPVAPRLIKNTES 277
QY 332 -----MPQ--LADRPFSFARICW-----CADTANREFLIDRHPQYHSLVLCGGA 373
DB 278 LEYGAHLVPEGLHSMFVQVAGNGWLLVAGDALRSCVNT-----GI 317

QY 374 SRGFKY----LPSIGNLIIVDAMEGKVPQKIHILIKWNPDIANRNRDTLGRFGG-PNR 428
DB 318 SVRGDMALTGAQAAQTLSACQHEPQNLPPL--YHNVERSLLM-DVLQRYQVHPAL 374
QY 429 VMDFDHVDKWTNVQYRDIDSKLGELEGLPIPNPLLRTGHHH 469
DB 375 LQRPGWYRTWPALM-ODISRLMDQDGKVPV-PLQLFWHH 413
RESULT 12
US-10-829-427-1
; Sequence 1, Application US/10829427
; Publication No. US20050026265A1
; GENERAL INFORMATION:
; APPLICANT: FURUKAWA, KEISUKE
; APPLICANT: KAJIYAMA, NAOKI
; TITLE OF INVENTION: MODIFIED SARCOSEINE OXIDASES, GENES AND RECOMBINANT DNAS THEREOF,
; FILE REFERENCE: 252202US0
; CURRENT APPLICATION NUMBER: US/10/829,427
; CURRENT FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: JP 2003-121533
; PRIOR FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: JP 2003-396807
; PRIOR FILING DATE: 2003-11-27
; PRIOR APPLICATION NUMBER: JP 2004-116345
; PRIOR FILING DATE: 2004-04-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-829-427-1

Query Match 6.6%; Score 166; DB 17; Length 387;
Best Local Similarity 21.4%; Pred. No. 5.4e-07;
Matches 89; Conservative 68; Mismatches 173; Indels 86; Gaps 20;
QY 20 LLIIVGAGTGTSTALHARCYNTVTVLPVPVPSAISAGNDVKNKVISGQYSNNKDEIE 79
DB 7 VIVVGAGSGMAAGYYLAKQGVKTLV-DAFDPHPHTEGSHHGDTIRREA-YGEGRKYP 64
QY 80 VNEILAEAFNGWKNDFLPKPYHDTGLL-----MSACSOEGL-----D 118
DB 65 F-ALRAQELWYLENETHNK-IFTKTGVLVFGPKGESDFVAETMEAAAEHSLTVDLEGD 122
QY 119 RLGVVRVPGEDPNLVELTRPEQFRKLAPEGVLQGDFFGWKGYFARSGAGNAHARNALVAA 178
DB 123 EINTR-WPG-----ITVPENYAI-----FEPNSGVLFENCIRSY 157
QY 179 AREAQRMGKVFVTGPQGRVVTLLIFENNVDKAVTGDGKIWRARTFLCAGASAGQFLDF 238
DB 158 RELAVAKGAKILTYT---RVEDFEVSQDVK-IQTANGS-YTADKLVSMGAWNSKLLSK 212
QY 239 KQQLRPTAWTLVHIALKPEERALYKNIPIVFN-----TERGFFPEPDERGEI-- 285
DB 213 LN-----LDIPLQYRVQVVGFFDSNEAKYSNDVGYPAFVMEVPKGIYYGFPSPFGG- 262
QY 286 IKICDBHPGVTNMQSADGTWMSIPFEKTOIPKEATRVALLKETMPQLADRPFSFARI 345
DB 263 ---CGLUKIGYHYTQOQIDPDTINREFGAYQ---EDESRLRDFLEKYMPE-ANGELKRGAV 315
QY 346 CWCADTANREFLIDRHPQYHSLVLCGASGRGFKYLPFSIGNLIIVD-AMEGKVPQKI 400
DB 316 CMYTKPTDHHFVTDTHPEHNSVFAAGFSGHGFKFSVSVVGEVLSQLATTGKTEHDI 371

RESULT 13
US-10-282-122A-68225

Sequence 68225, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zvekind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 68225
LENGTH: 385
TYPE: PRT
ORGANISM: Pseudomonas putida
US-10-282-122A-68225

Query Match 6.58; Score 165; DB 15; Length 385;
Best Local Similarity 23.5%; Pred. No. 6.7e-07;
Matches 101; Conservative 53; Mismatches 161; Indels 114; Gaps 22;

QY 16 KSSLLIVGAGTGTSTALHLARRGYNTVLDPPVPSPSAISAGNDVKNKVISS-----68
DB 6 KSYAVVVGUGVGSATLWFLAQ--QQDVL-----GLEGAPINLOGSSYGSRIF 55
QY 69 -GOYSNNKDBIEVNEILAEAFNGWK--NDPLFKPYVYDHTGLMSACSQBGLDRLGVRR 125
DB 56 ROAYWEGSDYLS---LLAEADLGWRELQATSHRPLLYHSG-----GLFIG 97
QY 126 PGSDPNLVELTRPEQFKLAPEGVLQDFFGKGYFARSGAGWAHARNALVAAREA---182
DB 98 P-----IRSGVWGSAAAS-----AKAG-GIAHOR--LTAAREVEARFS 131
QY 183 -----QRMGVKFVTGT-----PQGRVVTLLIPENNNDKGVAV---TG 214
DB 132 VFRADENMEAVFQGAFTIAADARLQMLNQAHAQMGFQSHVQDITRVESEFLLRLS 191
QY 215 DGKIMRAERTFLCAGAS-AGQFL-DFKNQLRPTAATLVHIALKPEERALKYKIPV-IFNI 271
DB 192 DGQSVLAQKVVLTATGAGLAGSLIPDLISGLLRPSRVPYIYWCFAKSGAEQLEAFNPFAYEL 251
QY 272 ERG--PFFPEDEERGE--IKICDEHPGYTNMVOSA-DGTMMSIPFEKTQIPKAEATRVRA 326

Db 252 EDGRLLYGTQIDNAPGIKI-----GFHNHQSQALDPRTOLEPASDAQIEE-----ISA 301
QY 327 LLKETMPQLADRPFPSPARICWCADTANREFLIDRHPQYSLVLCGASGRGFKYLPISGN 386
DB 302 CVSRVFPDLIARPYA-SRKCVYTWTPDEAFIIGESKELPSVFFVVSACSGHGFKAPALGS 360
QY 387 LIVDAMEGK 395
DB 361 CLARALAGQ 369

RESULT 14
US-10-156-761-9328
Sequence 9328, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 9328
LENGTH: 381
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-9328

Query Match 6.3%; Score 159.5; DB 14; Length 381;
Best Local Similarity 23.3%; Pred. No. 2.2e-06;
Matches 96; Conservative 53; Mismatches 172; Indels 91; Gaps 17;

QY 18 SLLIVGAGTGTSTALHLARRGYNTVLDPPVPSPSAISAGNDVKNKVISSGQYNNKDE 77
DB 5 ADVVWIGGVGVTSTIAYHLARAGVDVLRDELASGSTR-----AAGGVRAQPSDE 58
QY 78 IEVN-ELIAEEAFNGKNDPLFKPYVYDTC--LLMSACSQBGLDRLGVRRP--GEDPNL 132
DB 59 LNIRLGARSLAEAFRGFEPEPQDIGHRVGYLFLSTPEEVAAFETGVRLQNSLGVPSRM 118
QY 133 VELTRPEQFKLAP-----EGVLQGDFFGKGYFARSGAGWAHARNALVAAREAQRMGVK 188
DB 119 LE---PAEARRLSPLITDGLLAAAFSPDDGCHCTPESVHVGYA-----AAR---RHGAK 167
QY 189 F-----VTGTPOGRVVTLLIFENNNDKGVAVTGDGKIWAERTFLCA-----GASAGQF 235
DB 168 VLRHCEVTG-----IETRGDDITAVLRRGRI--GTGTVICAAGAWSKAVGAMVGD 217
QY 236 LDFKNQLRPTAATLVHIALKPEERALKYKIPVIFNTERGFFFPDEDEERGEIKICDEHPGY 295
DB 218 LPVRPURLRQIAVTEPVTVGLPP-----GLPMTIDFTSSLYFH-----TEGPGL 259
QY 296 TNNVQASDGTMSIPPEKTQIPKEATRVALLKETMPOLADRPFPSPARICWCADTANRE 355
DB 260 L-----VGMSDPDQTTGFTSTETHDRWIPRICEAMERRAPPLDLRR-----TGGWA 305
QY 356 FLIDRHPQYHSLV-----LGCASGRGFKYLPISGNLIVDAMEGKVP 397
DB 306 GLVEITPDHNLIGEAGSCSRFLYATGFGSHGFLQGPVAGVEVIRDLVLDVRP 357

RESULT 15

```

US-10-282-122A-55515
; Sequence 55515, Application US/10282122A
; Publication No. US20040029129A1
GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55515
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55515

```

Query Match	5.5%	Score 140;	DB 15;	Length 374;
Best Local Similarity	20.7%;	Pred. No. 0.00017;		
Matches	90;	Conservative 63;	Mismatches 147;	Indels 134; Gaps 22;
Qy	20	LLIVGAGTWGTSALHLARRGYTNVTVDPPVPVSATSAGDNVKNVTSSSGOYSNNKDEIE	79	
Db	7	LIITGSGVGAAGYATQAG-LKVLMTDAHLPPHSGSHGDRLLI---RHAYGEGERY	62	
Qy	80	VNEILLAEAFNGKNDPLFKP---YYHDTGLMSACSQ-----EGL	117	
Db	63	VPLVLRAQTL--W--DELAQTQEDRIERTGVINLGPASSEFLAGVERSAKAFNLDVEKL	118	
Qy	118	DLRGVRVSPGEDPNLVELTRPEQFRKL--APEGLVQGD--PFGKNGYVPARGAGWAHARN	173	
Db	119	DATGITAR-----WPEITVPDDYIGLEAHSGVLHCETAKTWIDLAAK--AGCAQLFN	170	
Qy	174	ALVAAAREAQRMGVKFVTGTTPQGRVVTLLIFENNDVKGAVTGDGKIWAERTFLCAGASAG	233	
Db	171	CPVEA-----ITHDADGVTVTTI-----DGG-YSASRLLVLSAG-----	202	
Qy	234	QLDFKNQLRPTAWT---LVHIALKPEERALLYKNIPVFNTERGFFFPDPDSERGEIKCD	290	
Db	203	-----TWYTRLLPELPQVPRK-----VFS-----WFQSD--GRYSAQN	234	

```

QY 291 EHPGYTNVQSADGTMSIPEKTCI-----PKEAETVRALLKETWEPQAD 337
Db 235 KPFAFTGELPND-QFYGFPSEKDKALKIKNGGQAIPTPEERN-----LFGAFPQDGS 287
QY 338 RPFESFAR-----ICWCADTANREFLIDRHQPYHSVLVLCGASGRGFKYLP SI 384
Db 288 EAFTEFLRTILPGVGLLYGAACYNTDPEDFIIDTLFGHDNVLLITGLSGHGFKFASVL 347
QY 385 GNLIVDAMEGKVPQ 398
Db 348 GEIAAQPAQGIAPQ 361

Search completed: May 27, 2005, 14:15:06
Job time : 138.87 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:28:55 ; Search time 103.659 Seconds
(without alignments)
2331.700 Million cell updates/sec

Title: US-10-622-893A-5

Perfect score: 2530

Sequence: 1 MGSGGDDDLALAVTKSSL.....LEGIPNPILRTGHHHHH 472

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2331	92.1	438	2 P78573	P78573 aspergillus
2	1982.5	78.4	437	2 Q92204	Q92204 aspergillus
3	1242.5	49.1	445	2 Q42629	Q42629 aspergillus
4	848	33.5	426	2 Q6BZB0	Q6BZB0 debaryomyce
5	846.5	33.5	418	2 Q6BZA7	Q6BZA7 debaryomyce
6	660	26.1	437	2 Q92261	Q92261 penicillium
7	629	24.9	438	2 Q96UT4	Q96UT4 emericella
8	627	24.8	437	2 Q765A9	Q765A9 eunicicilli
9	600	23.7	437	2 Q765A8	Q765A8 coniochaeta
10	488.5	19.3	412	2 Q43029	Q43029 schizosacch
11	465	18.4	433	2 Q9UTW9	Q9UTW9 schizosacch
12	426.5	16.9	502	2 Q7S638	Q7S638 neurospora
13	386	15.3	433	2 Q6CL10	Q6CL10 kluyveromyc
14	378.5	15.0	486	2 Q6BN43	Q6BN43 debaryomyce
15	351	13.9	446	2 Q6BN42	Q6BN42 debaryomyce
16	346	13.7	435	2 Q9P976	Q9P976 cylindrocarr
17	192	7.6	390	1 MSOX_BACB0	MSOX_BACB0 bacillus sp
18	185	7.3	387	2 Q6ITC6	Q6ITC6 bacillus ce
19	183	7.2	371	2 Q737A8	Q737A8 bacillus ce
20	181.5	7.2	440	2 Q9V9B5	Q9V9B5 drosophila
21	181	7.2	384	1 MSOX_STRAW	MSOX_STRAW streptomyce
22	181	7.2	388	1 MSOX_STRSB	MSOX_STRSB streptomyce
23	177.5	7.0	390	1 SOX_HUMAN	SOX_HUMAN homo sapien
24	177.5	7.0	390	2 Q6IJA9	Q6IJA9 homo sapien
25	175	6.9	371	2 Q63AM4	Q63AM4 bacillus ce
26	174	6.9	371	2 Q81PT2	Q81PT2 bacillus an
27	172.5	6.8	395	2 Q92NL2	Q92NL2 rhizobium m
28	172	6.8	367	2 Q982V8	Q982V8 rhizobium l
29	170.5	6.7	773	2 Q66IN9	Q66IN9 xenopus lae
30	170	6.7	371	2 Q81CN0	Q81CN0 bacillus ce
31	170	6.7	371	2 Q6H111	Q6H111 bacillus th

32	170	6.7	399	2	Q95U69	Q95U69 drosophila
33	169.5	6.7	390	1	SOX_RABIT	P79371 oryctolagus
34	169	6.7	435	2	Q8XTY8	Q8XTY8 raistonia s
35	166.5	6.6	423	1	YGCN_ECOLI	Q46904 escherichia
36	166	6.6	388	1	MSOX_ARTST	P40873 arthrobacte
37	165	6.5	382	2	Q88GE9	Q88GE9 pseudomonas
38	162.5	6.4	390	2	Q8N6Z6	Q8N6Z6 homo sapien
39	160	6.3	390	1	SOX_MOUSE	Q94826 mus musculu
40	159.5	6.3	381	2	Q82M71	Q82M71 streptomyce
41	159.5	6.3	410	2	Q89M92	Q89M92 bradyrhizob
42	157	6.2	433	2	Q8FEI5	Q8FEI5 escherichia
43	155	6.1	331	2	Q6R5L2	Q6R5L2 bacillus sp
44	155	6.1	375	2	Q8EMP0	Q8EMP0 oceanobacil
45	155	6.1	387	1	MSOX_BACSN	P23342 bacillus sp

ALIGNMENTS

RESULT 1

ID	P78573	PRELIMINARY;	PRT;	438 AA.
AC	P78573;			
DT	01-MAY-1997 (TrEMBLrel. 03, Created)			
DT	01-MAY-1997 (TrEMBLrel. 03, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Fructosyl amine: oxygen oxidoreductase.			
OS	Aspergillus fumigatus (Sartorya fumigata).			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			
OX	NCBI_TaxID=5085;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=37284723; PubMed=9139700; DOI=10.1074/jbc.272.19.12505;			
RA	Takahashi M., Pischetsrieder M., Monnier V.M.;			
RT	"Molecular cloning and expression of amadoriase isoenzyme (fructosyl			
RT	amine:oxygen oxidoreductase, EC 1.5.3) from Aspergillus fumigatus.";			
RL	J. Biol. Chem. 272:12505-12507(1997).			
DR	EMBL; U82830; AAC49711.1; -.			
DR	GO; GO:0003824; F:catalytic activity; IEA.			
DR	InterPro; IPR006076; Fad oxred.			
DR	InterPro; IPR000205; NAD_BS.			
DR	InterPro; IPR000594; Thif_domain.			
DR	Pfam; PF01266; DAO; 1.			
DR	SEQUENCE 438 AA; 48931 MW; F8C8A3ESD89D02AE CRC64;			

Query Match 92.1%; Score 2331; DB 2; Length 438;
Best Local Similarity 99.8%; Pred. No. 9.5e-171;
Matches 437; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	12	LAVTKSSLLIVGAGTGTSTALHLARRGTYNTVLDPPVPSAISAGNDVNVKVISGQY	71
Db	1	MAVTKSSLLIVGAGTGTSTALHLARRGTYNTVLDPPVPSAISAGNDVNVKVISGQY	60
Qy	72	SNKDEIENVNEILAEAFNCKNDPLPKPYHTDTGLMSACSQGLDRLGVVRPGEDPN	131
Db	61	SNKDEIENVNEILAEAFNCKNDPLPKPYHTDTGLMSACSQGLDRLGVVRPGEDPN	120
Qy	132	LVELTRPEQFRKLAPEGVLDQDPGWKGYFARSAGAWAHARNALVAAREAQRMGVKEVT	191
Db	121	LVELTRPEQFRKLAPEGVLDQDPGWKGYFARSAGAWAHARNALVAAREAQRMGVKEVT	180
Qy	192	GTQGRVVTLLIFENNVDKAVTGDGKIWRARTFLCAGASAGQFLDFKQNLRTPTAWTLVH	251
Db	181	GTQGRVVTLLIFENNVDKAVTGDGKIWRARTFLCAGASAGQFLDFKQNLRTPTAWTLVH	240
Qy	252	IALKPERALYKNIPIVFNTERGFFPEPDEBERGEIKICDEHPGTYNNVQSDGTMMSIPF	311
Db	241	IALKPERALYKNIPIVFNTERGFFPEPDEBERGEIKICDEHPGTYNNVQSDGTMMSIPF	300
Qy	312	EKTOIPKEASTRVALLKETMPQLADRFSPARCWADTANREFLIDRHPQYHSLVLGC	371
Db	301	EKTOIPKEASTRVALLKETMPQLADRFSPARCWADTANREFLIDRHPQYHSLVLGC	360

```
QY 372 GASGRGFKYLPSTGNLIVDAMEGKVPQKIHELKWNPDIAANRNWRDTLGRFGGPNRVMD 431
DB 361 GASGRGFKYLPSTGNLIVDAMEGKVPQKIHELKWNPDIAANRNWRDTLGRFGGPNRVMD 420

QY 432 FHDVKWNTVQYRDISKL 449
DB 421 FHDVKWNTVQYRDISKL 438

RESULT 2
Q92204 PRELIMINARY; PRT; 437 AA.
AC Q92204;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 23, Last annotation update)
DE Fructosyl amino acid oxidase.
GN Name=faaA;
OS Aspergillus terreus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_taxid=331178;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97175014; PubMed=9022674;
RA Yoshida N., Sakai Y., Isogai A., Fukuya H., Yagi M., Tani Y., Kato N.;
RT "Primary structures of fungal fructosyl amino acid oxidases important
RT for measurement of glycosylated proteins.";
RL Eur. J. Biochem. 242:499-505(1996).
DR EMBL; Y09020; CAA70218.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR PFam; PF01266; DAO; 1.
DR SEQUENCE 437 AA; 4869 MW; B83BEA0AB210536 CRC64;

Query Match 78.4%; Score 1982.5; DB 2; Length 437;
Best Local Similarity 82.2%; Pred. No. 6.1e-144;
Matches 360; Conservative 44; Mismatches 33; Indels 1; Gaps 1;

QY 12 LATVKSSLLIVGAGTGTSTALHLARGYTNVTLDPPVPSAISAGNDVNVKVISSGQY 71
DB 1 MPVTKSSLLIIAGTGTGCTSTALHLARGYTNVTLDPPVPSAISAGNDVNVKVISSGQY 60

QY 72 SNKDELEVNEILAEAFNGKNDPLPKPYHYDTGLLMSACSGEGLDRLGVVRPGEDPN 131
DB 61 SSKKDEVNEILAEQAFNGKNDPLPKPYHYDTGVVMSATVQGLERLGRVVRPEDEPD 120

QY 132 LVELTRPEQPKLAPGVLQDPPGKGYFARSAGWAHARNALVAAAREAQRMGVFT 191
DB 121 VAELTRPEQPKLAP-GVLKGNFPGRGYHRSNAGWAHARNALVAAAREAQRLGVRFVA 179

QY 192 GTPQGRVVTLLIFENNVDKAVTGDGKIWRABRTFLCAGASAGQFLDFKNQLRPTATLVH 251
DB 180 GSPQGRVITLLIFENNVDKAVTADGKIWRABRTFLCAGAAAGQFLDFKNQLRPTATLVH 239

QY 252 IALKPERALYKNIPVFNIERGFFPEPDEERGEIKICDEHPGYTNVQSDGTPMSIPF 311
DB 240 IQLKPERAQYKNMPVFNIEKGFFPEPDEERGEIKICDEHPGYTNVQSDGTRVMSIPF 299

QY 312 EKTQIPKEATRVALLKETMPQLADRPFSFARICWCADTANREFLIDRHPQYHSLVLC 371
DB 300 EKTQVPEAEWRVRLKSETMPQLADRPFSFARICWCADTANREFLIDRHPPEYPSLVLC 359

QY 372 GASGRGFKYLPSTGNLIVDAMEGKVPQKIHELKWNPDIAANRNWRDTLGRFGGPNRVMD 431
DB 360 GASGRGFKYLPSTGNLIVDAMEGKVPQKIHELKWNPDIAANRNWRDTLGRFGGPNRVMD 419

QY 432 FHDVKWNTVQYRDISKL 449
DB 420 FNEVKWNTVQYRDISKL 437

RESULT 3
```

```
042629 PRELIMINARY; PRT; 445 AA.
AC 042629;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fructosyl amine:oxygen oxidoreductase.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_taxid=5085;
RN [1]
RP SEQUENCE FROM N.A.
RA Takahashi M., Pischetsrieder M., Thein M., Roth J.C., Monnier V.M.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF035700; AAB88209.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR006076; Fad oxred.
DR InterPro; IPR000205; NAD_BS.
DR InterPro; IPR000594; Thif_domain.
DR PFam; PF01266; DAO; 1.
DR SEQUENCE 445 AA; 49364 MW; 57685E56CA535B23 CRC64;

Query Match 49.1%; Score 1242.5; DB 2; Length 445;
Best Local Similarity 51.1%; Pred. No. 5.3e-87;
Matches 231; Conservative 78; Mismatches 120; Indels 23; Gaps 7;

QY 14 VTKSSLLIVGAGTGTSTALHLARGYTNVTLDPPVPSAISAGNDVNVKVISSGQYSN 73
DB 6 LSTESSIIIVGAGTGTSTALHLARGYKDVTVLDPPHPVPSIAAGNDINKIM---EHSE 62

QY 74 NKD-----BIEVNEILAEAFNGKNDPLPKPYHYDTGLLMSACSGEGLDRLGV-RVRPG 127
DB 63 LKDGSDPRSAAFSTFTRALAKWKTDVPQYFHTGTGLISGHTALDIDHIDEVEPS 122

QY 128 EDPNLVELTRPEQPKLAPGVLQDPPGKGYFARSAGWAHARNALVAAAREAQRMGV 187
DB 123 E-TNFKLETAEDFRKTMPPGVLTGDPGKGLWHLKSGAGWIHAKKAMISAFNEAKRLGV 181

QY 188 KVTGTGTPQGRVVTLLIFENNVDKAVTGDGKIWRABRTFLCAGASAGQFLDFKNQLRPTAW 247
DB 182 RFVTGSPGPNVSLVYEDGDVGARTADGRVHKAHRTILSAGASDLSLDFKQLRPTAW 241

QY 248 TLVHIALKPEERALYKNIPVFNIERGFFPEPDEERGEIKICDEHPGYTNVQSD--GT 305
DB 242 TLCHIQMGPEEVKQYRNLPVLFNIAKGFPEPDEDEKHELKICDEHPGCYNFLPDPNRPQ 301

QY 306 MMSIPEKTKQIPKEATRVALLKETMPQLADRPFSFARICWCADTANREFLIDRHPQYH 365
DB 302 EKSVPFAKHQIPLEAEARARDFLHDTMPLADRPFSFARICWDADTPDRAFLIDRHPHP 361

QY 366 SLVILGCGASGRGFKYLPSTGNLIVDAMEGKVPQKIHELKWNPDIAANRNWRDTLGRFGG 425
DB 362 SLVAVVGGSGNGAMQMPETGGFIADALESKLQKVEKDIVWRPETAVDROWRATQNRFGG 421

QY 426 PNRVMDFHDVKE--WTNVQYRDISKLGLEGG 455
DB 422 PDRIMDFQVQGEDQWKI-----GESRG 444

RESULT 4
Q6BZB0 PRELIMINARY; PRT; 426 AA.
AC Q6BZB0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to tr|P78573 Aspergillus fumigatus Fructosyl amine:oxygen
DE oxidoreductase
DE ORFNames=DEHA0A03157g;
OS Debaryomyces hansenii CBS767.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.
```

```

OX NCBI_TaxID=284592;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confaniolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Sennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382133; CAG84414.1; -.
DR InterPro; IPR006076; Fad_oxred.
DR InterPro; IPR00205; NAD_BS.
DR Pfam; PF01266; DAO; 1.
SQ SEQUENCE 426 AA; 48297 MW; 7A5A382DF1F72293 CRC64;

Query Match      33.5%; Score 848; DB 2; Length 426;
Best Local Similarity 41.2%; Pred. No. 1.1e-56;
Matches 181; Conservative 71; Mismatches 153; Indels 34; Gaps 10;

QY 16 KSSSLIVAGTGTSTALHLARGYTNVTLPDPVPSAISAGNDVNKVISSGQ----Y 71
DB 3 KPGKILIGAGTGLSTALHLRQGEKDVLPDPVAVPSPFSAGNDVNKIIQTSSDDFF 62

QY 72 SNKKEIEVNEIIAEEAFNGKNDPLPKPYHYDHTGLLMSACSGEGLDRLGVRVR--PGED 129
DB 63 SK-----LALALELWREDNVFNKAFATGIYATGKEQESIDRYEYLLGRK 112

QY 130 PNLVELTRPQFRKLAP--EGV--LQGDFFGKGYFARSAGAGWAHARNALVAAREAQRM 185
DB 113 DKVVKLNSVDYEKYVPNKGKLSYPNKFQWGYQKNCGWAFARLALENCVCBECKL 172

QY 186 GVKFVTGTPQGRVTLIF--ENNDVKGAVTDGDKIWRARTFLCAGASAGOFDFKQQLRP 244
DB 173 GAKFVIDSAB-----ELFSDGACVGVHTSGNGNIIIEADRTIICAGANSFKFLNFEQQLA 228

QY 245 TMTVLVHIALKPERALYKNIPVFNIERGFFPEPDEERGEIKICDHPGYTNVQSDG 304
DB 229 KCTYLGHILKLTDEALLKGPVVLNDGGFVPEPD--LNNEIKFCNEFPGYVNVNED-- 285

QY 305 TWMISPEKTOIPKEAETRVRALLKETMPQLADRPFSFARICWCACTANREFLIDRHPOY 364
DB 286 ---SVPSFKDSIPKEAEDOMRFLRQVFPFAEREFSLARICWCTDTPDRHFLICEHPGH 342

QY 365 HSLVLGCGAGRGFKYLPSTGNLIVDM---EGKVPQKIHILKNWPDIAANRWRDTLG 421
DB 343 KNLVLGTDGSGQGFYKPNYNGKYSIQVALKGENSLDKDKELWRWRPDMGKRDLDKLOG 402

QY 422 RFGGNRVMDFHDVKEWTN 440
DB 403 RYGSNEVKDLKNVKQWSN 421

RESULT 5
Q6B2A7 PRELIMINARY; PRT; 418 AA.
ID Q6B2A7
AC Q6B2A7;

```

```

DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to tr|Q92204 Aspergillus terreus Fructosyl amino acid
DE oxidase.
GN ORFNames=DEHA0A032239;
OS Debaryomyces hansenii CBS767.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=284592;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confaniolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Sennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382133; CAG84414.1; -.
DR InterPro; IPR006076; Fad_oxred.
DR InterPro; IPR00205; NAD_BS.
DR Pfam; PF01266; DAO; 1.
SQ SEQUENCE 418 AA; 47714 MW; 448BCECF20CEB486A CRC64;

Query Match      33.5%; Score 846.5; DB 2; Length 418;
Best Local Similarity 42.0%; Pred. No. 1.4e-56;
Matches 181; Conservative 71; Mismatches 154; Indels 25; Gaps 10;

QY 16 KSSSLIVAGTGTSTALHLARGYTNVTLPDPVPSAISAGNDVNKVISSGQYNNK 75
DB 4 KESKILLIGGTGWLSTALYLLRDNYTDVIIAERIPSSRSAGNDINKIRIGYTSEFH 63

QY 76 DEIEVNEIIAEEAFNGKNDPLPKPYHYDHTGLLMSACSGEGLDRLGVRVRPGEDPN--L 132
DB 64 SRLEL-----EAAELWRNDPLKPFPHFVEGYLVCTSEAPLBELNIRIRLDNRKDW 117

QY 133 VELTRPQFRKLAPGVL--QGDFFGKGYFARSAGAGWAHARNALVAAREAQRMGVKFT 191
DB 118 EELTKEDIISGAP--ILHKGTLPGRGYLQREGGVHAANTWRAVANERIRLGGKFT 175

QY 192 GTPQGRVVTILFENNNDKGAVTGDGKIWRARTFLCAGASAGOFDFKQQLRPTAWTLVH 251
DB 176 ----DEISSLLFEAGSVKGCSTFTKEYPADKTLSTGASSVKLLDFKNQLMKWCWTLGH 231

QY 252 IALKPERALYKNIPVFNIERGFFPEPDEERGEIKICDHPGYTNVQSDGTMGMIPF 311
DB 232 IKLSTESERFKDIPVIRSLGEGFFQPD--ENNEMKICNEFFGYTHYISEKPEPNTSVPE 290

QY 312 EKTQIPKEAETRVRALLKETMPQLADRPFSFARICWCACTANREFLIDRHPOYH--SLAVG 370
DB 291 YINTVPVEAKGIRKLLKETIPDRDKPIVMSKICWCTDTPDRYLLIDEYSPESNGSLILA 350

QY 371 CGASGRGFKYLPSTGNLIVD--AMEGK--VPQKIHILKNWPDIAANRWRDTLGRFGGN 427
DB 351 TGDSGHGFQILPTIGKYISKLAVLGKGLDKEQOYRWRWPEITVKOR-----IQNRYGEG 406

```



```
QY 428 RVWDFHVKW 438
Db 407 HVKDLKIDW 417

RESULT 6
Q92261 PRELIMINARY; PRT; 437 AA.
AC Q92261;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fructosyl amino acid oxidase.
GN Names=faop;
OS Penicillium janthinellum (Penicillium vitale).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurociales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
OX NCBI_TaxID=5079;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97175014; PubMed=9022674;
RA Yoshida N., Sakai Y., Isogai A., Fukuya H., Yagi M., Tani Y., Kato N.;
RT "Primary structures of fungal fructosyl amino acid oxidases important
RT for measurement of glycosylated proteins.";
RL Eur. J. Biochem. 242:499-505(1996).
DR EMBL; Y09021; CAA70219.1; -.
DR Pfam; PF01266; DAO; 1.
SQ SEQUENCE 437 AA; 48231 MW; E681B8702DE4A574 CRC64;

Query Match 26.1%; Score 660; DB 2; Length 437;
Best Local Similarity 33.6%; Pred. No. 3.3e-42;
Matches 156; Conservative 75; Mismatches 157; Indels 76; Gaps 14;

QY 16 KSSLLIV-GAGTGTSTALHLARRGYT--NVTVDLPYPVPSAISAGNDVNVKVISSGQYS 72
Db 6 ESTKIVIVGGGTGSGSTALHLIRSGYTPSNITVLDVPIPSQAGYDLNKIMSI--RL 63

QY 73 NNKDEIVNEILAEAFNGKNDPLFKPYVYHDTGLMSACSQGLDRLGVVRVPGSDPNL 132
Db 64 RUGPDQ----LSLEALDMWKNDPLFKPFHNVGMIDVSSSTEEGIEGLRKYQSILLDAGILEKT 123

QY 133 -----VELTRPQFRLAPEGVLQDFFPGWKGYFARSGAGWAHARNALVAAAREAQRMG 186
Db 120 GLEKTNWLESEDDILAKAPH-FTREQIKGWKGLFCGDG-GWLAATAKAINAIGTFLKSQ 177

QY 187 VKFVTGT-----PQSRVVTILFENNVDKGTGDKIWRARTFLCAGASAGQFL 236
Db 178 VKFSGSAGTFKRLPAPDGTCS-----GVETVDGTYKVFADKVVLAAGAWSSTLV 228

QY 237 DFKNQLRPTATLVHIALKPEERALKYNIPIVFNIERGFFPEDEERGETIKIDEHPGYT 296
Db 229 DLEDQCVSKAWFAHQTLTPQESNQYKDVVVDGTYGFFPEPN-EHGVTKVCEFPGPS 287

QY 297 NM-----VQSGADGTMMSIPEKTQ-----IPKBAETRVALLKETMPQLADRPFSFARIC 346
Db 288 RFKLHPGYGATSPKLISVPSRSHAKPTDTPDSSEETIRKAIARFMPFKDKELFNRSKC 347

QY 347 WCADTANREFLIDRHPQYHSLVLCGASGRGFKYLPISIGNLIYDAMEGKVPQKIHILKW 406
Db 348 WCTDTADANLLICEHPKWNFILATGDSGHSFKVLPNIGKHVVLEIEGRLPQDLAGAWRW 407

QY 407 NP-----DIAANRNWRDTLGRFGGNRVMDFHDVK 436
Db 408 RFGDALKSKRSAPAKDLAEMPWK-----HDAK 436

RESULT 7
Q96UT4 PRELIMINARY; PRT; 438 AA.
AC Q96UT4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fructosyl peptide oxidase (EC 1.5.3.-).

DE Fructosyl amino acid oxidase.
GN Name=faoa;
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurociales; Trichocomaceae; Emericella.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=FGSC A4;
RX MEDLINE=22261769; PubMed=12375102;
RA Jeong H.Y., Song M.H., Back J.H., Han D.M., Wu X., Monnier V.,
RA Jahng K.Y., Chae K.S.;
RT "The veA gene is necessary for the inducible expression by fructosyl
RT amines of the Aspergillus nidulans faoA gene encoding fructosyl amino
RT acid oxidase (amadoriase, EC 1.5.3).";
RL Arch. Microbiol. 178:344-350(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=FGSC A4;
RA Jeong H.-Y., Song M.H., Back J.H., Han D.M., Monnier V.M.,
RA Jahng K.-Y., Chae K.-S.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF415568; AAL09028.1; -.
DR InterPro; IPR006076; Fad_Oxred.
DR InterPro; IPR000205; NAD_BS.
DR Pfam; PF01266; DAO; 1.
SQ SEQUENCE 438 AA; 48433 MW; BF72BC2CC4D57316 CRC64;

Query Match 24.9%; Score 629; DB 2; Length 438;
Best Local Similarity 34.0%; Pred. No. 8.1e-40;
Matches 148; Conservative 73; Mismatches 168; Indels 46; Gaps 12;

QY 20 LLIVAGTGTGTSTALHLARRGYT--NVTVDLPYPVPSAISAGNDVNVKVISSGQYSNNKDE 77
Db 10 IVGGGTGSGSTALHLRAGYTPSNITVLDTCPIPSAQSGYDLNKIMSI--RLNKPD 67

QY 78 IEVNEILAEAFNGKNDPLFKPYVYHDTGLMSACSQGLDRLGVVRP-----GEDPN 131
Db 68 LQ-----LSLEALDMWKNDPLFKPFHNVGMIDVSSSTEEGIEGLRKYQSILLDAGILEKT 123

QY 132 LVELTRPQFRLAPEGVLQDFFPGWKGYFARSGAGWAHARNALVAAAREAQRMGKVFVT 191
Db 124 NFMLESEDDILAKAPH-FTREQIKGWKGLFCGDG-GWLAATAKAINAIGTFLKEQGVKFGF 181

QY 192 GTPQGRVVTILFENNVDK-----GAVTGDGKIWRARTFLCAGASAGQFLDKQLRPTAT 248
Db 182 G-GAGTFKXPLFADAHEKTCIGVEITVDGTYKVFADKVVLAAGAWSSTLVLEECVSKAW 240

QY 249 LVHIALKPEERALKYNIPIVFNIERGFFPEDEERGETIKIDEHPGYTMVQ-----SAD 303
Db 241 FAHIQLTPAAEAAYKNTPIVDGTYGFFPEPN-ENGIIKVCDFPGTFHFKMHQPYGSPA 299

QY 304 GTWMSITPPEKTQIPKE-----AETRVALLKETMPQLADRPFSFARICWCADTANREFLI 358
Db 300 PKPISVPSRSHAKPTDTPHASEVTIKKAINRPLPRFNDKELFNRAWCWCTDTADANLLV 359

QY 359 DRHPQYHSLVLCGASGRGFKYLPISIGNLIYDAMEGKVPQKIHILKNP----- 408
Db 360 CEHPRWKGYLATGDSGHSFKLLPNIGKHVVLEIEERLESVFKDANWRPFGSGDALKSR 419

QY 409 -----DIAANRNWRD 418
Db 420 AAPAKDLADMPGWRN 434

RESULT 8
Q765A9 PRELIMINARY; PRT; 437 AA.
AC Q765A9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Fructosyl peptide oxidase (EC 1.5.3.-).
```


[illegible][illegible]

```

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schafer M., Muller-Auer S.,
RA Gabel C., Fuchs M., Dusterhoft A., Fritz C., Holzer E., Moestl D.,
RA Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R.,
RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
DR EMBL; AL022071; CAA17815.1; -.
DR PIR; T40295; T40295.
DR GeneDB SPombe; SPBC354.15; -.
DR InterPro; IPR006076; Fad_oxred.
DR InterPro; IPR00205; NAD_BS.
DR Pfam; PF01266; DAO; 1.
SQ SEQUENCE 412 AA; 46319 MW; BD36B86075D3F4EF CRC64;

Query Match 19.3%; Score 488.5; DB 2; Length 412;
Best Local Similarity 30.1%; Pred. No. 4.8e-29;
Matches 130; Conservative 78; Mismatches 183; Indels 41; Gaps 12;

QY 14 VTKSSLLIVGAGTGTSTALHARGLYNTVTLDPVPVPSAISAGNDVNVKISSGOYSN 73
DB 1 MVKNTSVIIVGAGVGLSAALETKGGYTIKILDRAPPVIDGSSVDANRIIS-DYAD 59

QY 74 NKDEIEVNEILAEAFNGMKNDPLFKPYHYHDTGLMSA-----CSQGLDRLGVYVR 125
DB 60 -----AVYCSMGIDALEEWRNPLFKQFYGSLMFVGRDNRVYRDMSLNLTGMGV--- 111

QY 126 PGEDPNLVELTRQFRKLAPEGLVQDFFGWKGYFARSAGAGAHARNALVAAAREAQM 185
DB 112 -----SAAKFQTTLEUKLGFPPKI--GELNDGAGYANFSSGWANAEPQSVKSVVNYLAH 164

QY 186 GVKFVTGTPOGRVTVLIFENNVDKAGTGDGKTIWRAERTPLCAGASAGQFL-DFKNQLRP 244
DB 165 GVSVFISG-PGTYVBELITENVVGVRTTGTG-AYMAEKLIPTAGTAWTASLLPNDHTFLA 222

QY 245 TAWTLVHIALKPEERALKYKNIPVFNIERGFFPEPDEERGEIKICDEHPGYTNVQSADG 304
DB 223 TGGPVAYIKLTPEYIRFLTNPNVLPDFTGTYIFPPPTPDGKLFARHGYGFTRMQNLKSG 282

QY 305 TWMISPEKTIQI-----PKEATRVALLKETM-PQLADRPFSFARICWCADTANREFLD 359
DB 283 KVESVPKPKPLVLPKAEALDLRLNLTQTYGEEISQRFYKTRICYDITDADAEVFD 342

QY 360 RHPQYHSLVLCGASGRGKYLPSIGNLIVDAMEGKVPQKIHELKWK---NPDIARNNW 416
DB 343 YHPDYENLVCTGGSGHGFFAFLPILGKYSIGCMFRELLEPLKKRWKKNLENLEFALDHS 402

QY 417 RDTLGRFGGPNR 428
DB 403 R-----AGFSR 408

RESULT 11
Q9UTM9 PRELIMINARY; PRT; 433 AA.
AC Q9UTM9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SPAC139.04c protein.

```

```

GN Name=SPAC139.04c;
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Spouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris T., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schafer M., Muller-Auer S.,
RA Gabel C., Fuchs M., Dusterhoft A., Fritz C., Holzer E., Moestl D.,
RA Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R.,
RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA McDougall R.C., Rajandream M.A., Barrell B.G., Saunders D., Harris D.;
RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL132667; CAB59618.1; -.
DR PIR; T37605; T37605.
DR GeneDB SPombe; SPAC139.04c; -.
DR InterPro; IPR006076; Fad_oxred.
DR InterPro; IPR000583; GATase_2.
DR Pfam; PF01266; DAO; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; UNKNOWN 1.
SQ SEQUENCE 433 AA; 48973 MW; 170081BB3D48B1BE CRC64;

Query Match 18.4%; Score 465; DB 2; Length 433;
Best Local Similarity 30.1%; Pred. No. 3.3e-27;
Matches 130; Conservative 80; Mismatches 174; Indels 48; Gaps 13;

QY 17 SSSLIVGAGTGTSTALHAR-RGYNTVTLDPVPVPSAISAGNDVNVKISSGOYSNNK 75
DB 2 SRTIVIVGCVGLSTAVELAKNHSFDNIIDAEPVPVSSMSAANDINKIVRP-EYADLK 60

QY 76 DELEVNEILAEAFNGMKNDPLFKPYHYHDTGLMSACSQGLDRLGVYVRPGE----- 128
DB 61 -----YMKLALEAMEKWRNDPELSSVYFECGRUST-----ISKPYRARFDEVAQRNL 109

QY 129 ----DPNLVELTRQFRKLAPEGLVQDFFGWKGYFARSAGAGAHARNALVAAAR 180
DB 110 KLGLDSALINLSSEIRKYPFLSNSPLRSDMQA-----VNVHAGYANSAASKLLEL 165

QY 181 EAQRMGVKFTGTPQGR-----VV-----TLFENNVDKAGTGDGKTIWRAERTPLCAGASA 232
DB 166 KARELGVEFVFG-KAGKFKFVNVHSETDIDKNDHVSQVQTEDGTIYHADTILLAVGAYL 224

QY 233 GQFLDFKNQLRPTAWTLVHIALKPEERALKYKNIPVFNIERGFFPEPDEERGEIKICDEH 292
DB 225 NAYLNTSHRVCAKGLPVAHIQLTDBEFKTYKNMPLIIFDPDCAAFVPPYPTVKLIKLASTG 284

```

```

QY 293 PGYTNMQS---ADGTMMSIPP---EKTQIPKEATRVALLKTMQADRPSPFARIC 346
DQ 285 YEYVCNVETDENSQVSVPHSGPSKSLPKYAIQWRFLDTFLPLADRSINTKWC 344
QY 347 WCADTANREFLIDRHPQVSHSLVLCGASGRGFKYLPSTGNLIVDAMEGKVPQKTHELIKW 406
DQ 345 WISTEDANFLIDKVPQFDNVFVANGDSGAFKFLPNIGRYIAQRIILGDISSEWKDAWRW 404
QY 407 NPDTAANR-NWR 417
DQ 405 REDDKASELKWR 416

RESULT 12
QY 7S638 PRELIMINARY; PRT; 502 AA.
AC Q7S638;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 25-OCT-2004 (TrEMBLrel. 26, Last sequence update)
DE Hypothetical protein (Related to fructosyl amino acid oxidase).
GN Name=NCU04771.1; Synonyms=B15B10.150;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Furtell S., Rehman B.,
RA Ekins T., Engels S., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysheles M., Maucelli E., Bielke C., Rudd S., Frishman D.,
RA Kryzofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Varden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbold D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000331; EAA310984.1; -.
DR EMBL; BX897677; CAE85581.1; -.
DR InterPro; IPR006076; FAD_oxred.
DR Pfam; PF01266; DAO; 1.
KW Hypothetical protein.
SQ SEQUENCE 502 AA; 52930 MW; 39748F9E553D4FDE CRC64;

```

```

Query Match      16.98; Score 426.5; DB 2; Length 502;
Best Local Similarity 29.24; Pred. No. 3.7e-24;
Matches 138; Conservative 56; Mismatches 169; Indels 109; Gaps 21;

QY 18 SSLIVGAGTGWGTSALHLARR-GYTNVTVL-----DPYFVPSAISAGNDVNKVI 66
DQ 30 SSLILIGSGVGLTAYALSRHDFVNTTITVVDADPSPDPVFPAPDAASV--DTSRII 87

```

```

QY 67 SSGOYNNKDEIEVNEILABEAFNGWK-----NDPLPKPYHYDITGLMSACSQSG-LDRL 120
DQ 88 RA-DYPD-----AAYAALAAEAQLQWRQTHDPDLGSEGRYSSEGLLVVA---DGPAPPP 138
QY 121 GVRVRPG-----EDPNLV-----ELTRPEQFRKLAPGV 149
DQ 139 GTPVAPGTSVIDKSLTGM DYARPSWANVLSSASDPELAARIKECPNTDAIKALGTG- 197
QY 150 LQGGDFPGWKGYFARSGAGWAHARNALVAAAREAQMG-VKFPVTGTPOGRVVTLLIFENNDV 208
DQ 198 --GSSGSW-GYI-NGLSGWANAGASMAWLVKRVRAEGRINFVA---GEVTNLEVSNTV 249
QY 209 KGAVTGDKIWRARTFLCAGASAGQFLDFKNQLRPTATWLHVIALKPEERALYKNIPVI 268
DQ 250 TGAFLSDGRVLSADLVNMSAGAWTGRVLDTGQATGQVGLGYIDLTPEEEAQLAHMPVI 309
QY 269 FNTERGFPFDEERGEIKICDEHPGYTNMVQSADGTMMSP-----PEKT----- 314
DQ 310 LNLSTGLFVIP-PRNGVLKVARHAYGYLN-----PTLSVPPLASFTTTPAAAVSLPLT 362
QY 315 -----QIPKEATRVALLKE--TMPQLADRPSPFARICWCADTANREFLIDRHPQYH 365
DQ 363 TLTDPTLQIPTGAGADLRRALHEMVPLSLRDRPFSTRICWCYSDTPTADPIVDYHPKYK 422
QY 366 SLVLGCGASGRGPKYLPSTGNLIVDAMEGKVPQKIHILKWNPDIAANRWR 417
DQ 423 GLFVATGDSGHAFKFLPVIKEIADVIAGQCP-----PREFVGKWNMR 464

```

RESULT 13

```

QY 6Q6CLIO PRELIMINARY; PRT; 433 AA.
AC Q6CLIO;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to sp|Q9UTW9 Schizosaccharomyces pombe Putative fructosyl
DE amino acid oxidase.
GN ORFNames=KLIA0F028609;
OS Kluyveromyces lactis NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talia E.,
RA Goffard N., Frangeul L., Aigle M., Anthonard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boistame A., Boyer J., Cattolico L., Confaniolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerret A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikoleki M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellens S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RA Genoscope;
RL Submitted
DR EMBL; CR382126; CAG97917.1; -.
DR InterPro; IPR006076; FAD_oxred.
DR InterPro; IPR000205; NAD_BS.
DR Pfam; PF01266; DAO; 1.
SQ SEQUENCE 433 AA; 48905 MW; 1091A3C6C3F3DA08 CRC64;

```


Search completed: May 27, 2005, 13:47:31
Job time : 104.659 secs

THIS PAGE BLANK (USPTO)

Result	No.	Query			DB	ID	Description
		Score	Match	Length			
1	1982.5	78.4	437	2	AAW39253	AAw39253 A. terreu	
2	1982.5	78.4	437	2	AAW29100	Aay29100 A. terreu	
3	1977.5	78.2	437	2	AAW37141	Aaw37141 Aspergill	
4	1158.5	45.8	440	5	AAW69251	Aaw69251 Fructosyl	
5	1158.5	45.8	440	5	ABG30782	Abg30782 Fusarium	
6	1152.5	45.6	440	5	ABG30784	Abg30784 Fusarium	
7	1151.5	45.5	440	5	ABG30785	Abg30785 Fusarium	
8	1150.5	45.5	440	5	ABG30783	Abg30783 Fusarium	
9	1115.5	44.1	444	8	Adm78789	Adm78789 Fusarium	
10	837.5	33.1	427	8	ADSL5771	Adsl5771 Fructosyl	
11	661.26	26.1	432	7	ADP81213	Adp81213 Orf20, SE	
12	660.26	26.1	437	2	AAW24134	Aaw24134 Fructosyl	
13	660.26	26.1	437	2	AAW97386	Aaw97386 A fructos	
14	660.26	26.1	437	2	AAW97387	Aaw97387 A fructos	
15	627.24	24.8	437	6	ABR41990	Abra41990 Eupenicil	
16	600.23	23.7	437	6	ABR41989	Abra41989 Coniocha	
17	437.5	17.3	439	5	ABM78358	Abm78358 Amino aci	
18	416.5	16.5	473	8	ADM78787	Adm78787 Fusarium	
19	330.5	13.1	109	2	AAW69252	Aaw69252 Fructosyl	
20	198	7.8	411	7	ADC59522	Adc59522 Arthrobas	
21	192	7.6	390	1	AAW94663	Aap94663 New sarco	
22	192	7.6	390	2	AAW07397	Aaw07397 Bacillus	
23	184	7.3	389	2	AAW23780	Aar23780 Sarcosine	
24	181.5	7.2	440	4	ABW59342	Abw59342 Drosophil	
25	177.5	7.0	390	4	ABR08949	Abbr08949 Human ser	

CC acting on glycoamino acids. This sequence represents a polypeptide used
 CC in the scope of the invention
 XX
 SQ Sequence 440 AA;

Query Match 45.6%; Score 1152.5; DB 5; Length 440;
 Best Local Similarity 50.9%; Pred. No. 3.6e-105;
 Matches 223; Conservative 75; Mismatches 133; Indels 7; Gaps 5;

Qy 11 ALAVTKSSLLIVAGTGTSTALHLARRGYTNVTLDPYVPSAISAGNDVNVKVIS--S 68
 Db 1 ASTLTQSQILIVGGGTWGCSTALHLARRGYTNVTLDPYVPSAISAGNDVNVKLAGRLS 60

Qy 69 GOYSNNKDEIEVNEILAEAFNGKNDPLFKPYHTDTGLLMSACSQGLDLRGVVRPGE 128
 Db 61 TADSKGDEDSIWKALSAAAQGLHDPVFPFCHNTGTVVAGSTPKSIKQL-VEDEIGD 119

Qy 129 D-PNLVELTRPEQFRKLAPGVLQDPPGKGYFARSAGAWAHARNALVAAAREARMGV 187
 Db 120 DIDQYTPLTAEPRKTMPEGILTGNPFGKGYKPTGSGVHARKAKAAFESEERLGV 179

Qy 188 KFTVTGPGQVRVTLIFENNNDVKGAVTDGDKIWRARTEFLCAGASAGQFLDKQLRPTAW 247
 Db 180 KFTTSGPEGKVESLIFEDGVRGAKTADGKEHRADRTILSAGASAEFFLDFENQIQPTAW 239

Qy 248 TLVHIALKPERALYKNIPIVNIERGEFFPEDEERGEIKICDEHPGYTNVWQSDGTM- 306
 Db 240 TLGHIQITPEETKLYKNLPFLFNQGFMEPEDLHQLKMCDEHPGVCNWEKPGSKYP 299

Qy 307 MSIPPEKTOIPKEAETRVRLAKKETMPQLADRPFSFARICWCADTANREFLIDRHPYHS 366
 Db 300 QSIPTAKHQVPTAEARMKQFLKDIIMPQLADRLVHARICWCADTQDRMFLITYHPHPS 359

Qy 367 LVLGCGASGRGFKYLSIGNLIVDAMEGKVPQKIHILIKWNPDIANRNRDITLGRFGGP 426
 Db 360 LVIASGDCGTYVHITSIGKFIISDCMEGTLEERFAKYWRWRPEKFTFEGWKDPLDRFGAD 419

Qy 427 NRWVDF--HDVKEWTNVQ 442
 Db 420 DKIMDLPKSDVEGWNTIK 437

RESULT 7
 ABG30785
 ID ABG30785 standard; protein; 440 AA.
 XX
 AC ABG30785;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Fusarium oxysporum polypeptide #4.
 XX
 KW Glycoamino acid; glycoprotein; deoxycholic acid; deoxycholic amide;
 KW cholic acid amide octyl glucoside; quaternary ammonium salt;
 KW quaternary ammonium salt cationic surfactant; concanavalin A; betaine;
 KW ascorbic oxidase; 4-(2-hydroxyethyl)-1-piperazinyl group; albumin;
 KW protease; glycoalbumin; bromocresol purple; glycation; diabetes;
 KW ascorbic acid; mutant; mutein.
 XX
 OS Fusarium oxysporum.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 246
 FT /note= "Encoded by ATG"
 FT Misc-difference 372
 FT /note= "Wild-type Lys substituted by Val"
 FT Misc-difference 396
 FT /note= "Encoded by TTC"
 XX
 PN WO200261119-A1.
 XX
 PD 08-AUG-2002.

XX
 PF 30-JAN-2002; 2002WO-JP000721.
 XX
 PR 31-JAN-2001; 2001JP-00022953.
 PR 16-FEB-2001; 2001JP-00037996.
 PR 08-AUG-2001; 2001JP-00240002.
 XX
 PA (ASAH) ASAH KASEI KOGYO KK.
 XX
 PI Kouzuma T, Yoshioka I, Arai M, Sumitani J, Imamura S;
 XX
 XX WPI; 2002-599854/64.
 DR N-PSDB; ABK30478.
 DR
 XX
 PT Compositions for assaying glycoprotein without interference from globulin
 PT and ascorbic acid in blood, useful in clinical examination including
 PT diagnosis and management of diabetes.
 XX
 PS Disclosure; Page 72-75; 82pp; Japanese.
 XX
 CC The invention relates to compositions for assaying glycoprotein with use
 CC of protease and enzymes acting on glycoamino acids in the presence of at
 CC least 1 of deoxycholic acid or amide, cholic acid amide octyl glucoside,
 CC quaternary ammonium salt, quaternary ammonium salt cationic surfactant,
 CC concanavalin A and betaine and/or ascorbic oxidase and a buffer without a
 CC 4-(2-hydroxyethyl)-1-piperazinyl group. The invention also relates to
 CC compositions for assaying albumin containing the protease and enzymes for
 CC determining glycoalbumin and a protein-denaturing agent and/or a compound
 CC having S-S bond and bromocresol purple for evaluation of glycation
 CC proportion of albumin into glycoalbumin with respect to the other
 CC albumins measured separately. The compositions are useful in clinical
 CC examination including diagnosis and management of diabetes. The
 CC compositions can be used in methods which avoid the effects of globulin
 CC and ascorbic acid components and stabilise enzymes such as protease
 CC acting on glycoamino acids. This sequence represents a polypeptide used
 CC in the scope of the invention
 XX
 SQ Sequence 440 AA;

Query Match 45.5%; Score 1151.5; DB 5; Length 440;
 Best Local Similarity 50.9%; Pred. No. 4.5e-105;
 Matches 223; Conservative 75; Mismatches 133; Indels 7; Gaps 5;

Qy 11 ALAVTKSSLLIVAGTGTSTALHLARRGYTNVTLDPYVPSAISAGNDVNVKVIS--S 68
 Db 1 ASTLTQSQILIVGGGTWGCSTALHLARRGYTNVTLDPYVPSAISAGNDVNVKLAGRLS 60

Qy 69 GOYSNNKDEIEVNEILAEAFNGKNDPLFKPYHTDTGLLMSACSQGLDLRGVVRPGE 128
 Db 61 TADSKGDEDSIWKALSAAAQGLHDPVFPFCHNTGTVVAGSTPKSIKQL-VEDEIGD 119

Qy 129 D-PNLVELTRPEQFRKLAPGVLQDPPGKGYFARSAGAWAHARNALVAAAREARMGV 187
 Db 120 DIDQYTPLTAEPRKTMPEGILTGNPFGKGYKPTGSGVHARKAKAAFESEERLGV 179

Qy 188 KFTVTGPGQVRVTLIFENNNDVKGAVTDGDKIWRARTEFLCAGASAGQFLDKQLRPTAW 247
 Db 180 KFTTSGPEGKVESLIFEDGVRGAKTADGKEHRADRTILSAGASAEFFLDFENQIQPTAW 239

Qy 248 TLVHIALKPERALYKNIPIVNIERGEFFPEDEERGEIKICDEHPGYTNVWQSDGTM- 306
 Db 240 TLGHIQITPEETKLYKNLPFLFNQGFMEPEDLHQLKMCDEHPGVCNWEKPGSKYP 299

Qy 307 MSIPPEKTOIPKEAETRVRLAKKETMPQLADRPFSFARICWCADTANREFLIDRHPYHS 366
 Db 300 QSIPTAKHQVPTAEARMKQFLKDIIMPQLADRLVHARICWCADTQDRMFLITYHPHPS 359

Qy 367 LVLGCGASGRGFKYLSIGNLIVDAMEGKVPQKIHILIKWNPDIANRNRDITLGRFGGP 426
 Db 360 LVIASGDCGTYVHITSIGKFIISDCMEGTLEERFAKYWRWRPEKFTFEGWKDPLDRFGAD 419

Qy 427 NRWVDF--HDVKEWTNVQ 442
 Db 420 DKIMDLPKSDVEGWNTIK 437

Db 420 DKIMDLPKSDVEGWNIK 437

RESULT 8

ABG30783

ID ABG30783 standard; protein; 440 AA.

AC ABG30783;

XX

DT 05-NOV-2002 (first entry)

XX

DE Fusarium oxysporum polypeptide #2.

XX

KW Glycoamino acid; glycoprotein; deoxycholic acid; deoxycholic amide;

KW cholic acid amide octyl glucoside; quaternary ammonium salt;

KW quaternary ammonium salt cationic surfactant; concanavalin A; betaine;

KW ascorbic oxidase; 4-(2-hydroxyethyl)-1-piperaziny group; albumin;

KW protease; glycoalbumin; bromocresol purple; glycation; diabetes;

KW ascorbic acid; mutant; mutein.

XX

OS Fusarium oxysporum.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Misc-difference 246 /note= "Encoded by ATG"

FT Misc-difference 372 /note= "Wild-type Lys substituted by Trp. Encoded by TTC"

FT Misc-difference 396 /note= "Encoded by TTC"

FT

XX

PN W0200261119-A1.

XX

PD 08-AUG-2002.

XX

PF 30-JAN-2002; 2002WO-JP000721.

XX

PR 31-JAN-2001; 2001JP-00022953.

PR 16-FEB-2001; 2001JP-00039796.

PR 08-AUG-2001; 2001JP-00240002.

XX

PA (ASAH) ASAH KASEI KOGYO KK.

XX

PI Kouzuma T, Yoshioka I, Arai M, Sumitani J, Imamura S;

XX

DR WPI; 2002-599854/64.

DR N-PSDB; ABK90476.

XX

XX Compositions for assaying glycoprotein without interference from globulin and ascorbic acid in blood, useful in clinical examination including diagnosis and management of diabetes.

PS Disclosure; Page 66-69; 82pp; Japanese.

XX

CC The invention relates to compositions for assaying glycoprotein with use of protease and enzymes acting on glycoamino acids in the presence of at least 1 of deoxycholic acid or amide, cholic acid amide octyl glucoside, quaternary ammonium salt, quaternary ammonium salt cationic surfactant, concanavalin A and betaine and/or ascorbic oxidase and a buffer without a 4-(2-hydroxyethyl)-1-piperaziny group. The invention also relates to compositions for assaying albumin containing the protease and enzymes for determining glycoalbumin and a protein-denaturing agent and/or a compound having S-S bond and bromocresol purple for evaluation of glycation proportion of albumin into glycoalbumin with respect to the other albumins measured separately. The compositions are useful in clinical examination including diagnosis and management of diabetes. The compositions can be used in methods which avoid the effects of globulin and ascorbic acid components and stabilise enzymes such as protease acting on glycoamino acids. This sequence represents a polypeptide used in the scope of the invention

XX

SQ Sequence 440 AA;

Query Match 45.5%; Score 1150.5; DB 5; Length 440;

Best Local Similarity 50.9%; Pred. No. 5.7e-105;

Matches 223; Conservative 75; Mismatches 133; Indels 7; Gaps 5;

QY 11 ALAVTKSSLLIVGAGTGTSTALHLARRGYTNVTVDLPYPVPSAISAGNDVNVKVIS--S 68

DB 1 ASTLTQSQQLIVGGTGWGCTALHLARRGYTNVTVDLVDNRIPSPISAGHDVNVKLAGRLS 60

QY 69 GOYSNNKDEIEVNEJLAEEAFNGKNDPLFKPYHYDHTGLLMSACSOEGDLRLGVRVRPGE 128

DB 61 TADSKGDDSDSIWKALSYAAQGLHDPVFPQFCHNTGVSVAGSTPKSIKQL-VEDEIGD 119

QY 129 D-PNLVELTRPEQFRKLAPEGLVQDFFQGWKGYFARSAGWAGHARNAALVAAREAGRMGV 187

DB 120 DIDQVTPLTAAEDFRKTMPEGILTGNFPQGWKGYFKPTGSGWVHARKAMKAFAEESRLGV 179

QY 188 KFTVTGPQGRVTLTFENNVDKGVATGDGKIWRABRTFLCAGASAGQFDFKNQLRPTAW 247

DB 180 KFTTSGPEGKVESLTFEDGDVGRGAKTADGKEHARDRTILSAGASAEFFLDFENQIQPTAW 239

QY 248 TLVHIALKPEERALYKNIPIVFNIERGFFFEDEERGEIKICDEHPGYTNMVQSADGTM- 306

DB 240 TLGHQITPEETKLYKNLPFLFNINQGFWEDEDLHQLKMCDEHGYCNWVEKPSKYP 299

QY 307 MSIPPEKTOIPKEAETRVALLKETNPQLADRPFSFARICWCADTANRBFIDRHPQYHS 366

DB 300 QSIPFAKHQVPTAEARRMKQFLKDIMPQLADRPVHARICWCADTQDRMFLITYHPRHS 359

QY 367 LVLGCGASGRGPKYLPISIGNLIVDAMEGKVPQKIHELKKNPDIAANRNWRDTLGRFGGP 426

DB 360 LVIASGDCGTGYWHITSIGKFISSDCMGTLEERFAKYWRWRPEKFTFEGKQDPLDRFGAD 419

QY 427 NRVMDF--HDVKEWTNVQ 442

DB 420 DKIMDLPKSDVEGWNIK 437

RESULT 9

ADM78789

ID ADM78789 standard; protein; 444 AA.

XX

AC ADM78789;

XX

DT 01-JUL-2004 (first entry)

XX

DE Fusarium proliferatum fructosylamine oxidase Q2 protein sequence.

XX

KW fructosylamine oxidase; fructosyllysine; fructosylvaline;

KW Amadori compound; medical examination; diagnosis; serum glucose level;

KW food inspection; enzyme.

XX

OS Fusarium proliferatum.

XX

PN W02004029251-A1.

XX

PD 08-APR-2004.

XX

PF 16-SEP-2003; 2003WO-JP011766.

XX

PR 24-SEP-2002; 2002JP-00277214.

PR 24-OCT-2002; 2002JP-00309734.

XX

PA (ARKR-) ARKRAY INC.

XX

PI Yoshida N, Tani Y, Yonehara S;

XX

DR WPI; 2004-316127/29.

DR N-PSDB; ADM78788.

XX

PT Fusarium proliferatum-originated fructosylamine oxidase useful in measurement of Amadori compound and particularly applicable in medical examination including diagnosis and management of diabetes.

XX

Db	288	RFKLHQPVGATSPKLI	SVPRSHAKHPTDTP	DDSSSETIRKAIARFWPRFKDKELFNRSWC	344				
Qy	347	WCADTANREFLIDRHPQYHSLV	LGCSAGRGFKYLP	SLIGNLIIVDAMEGKVPQKIHLEIKW	406				
Db	348	WCTDTADANLLICEHPKWKNFILAT	GDGSHGSKVL	PNIGKHVVELIEGRLPQDLAGAWRW	407				
Qy	407	NP-----	DIAANRNWRD	TLGRFGGPNRVMDFDHVK	436				
Db	408	RPGGDALKRSAPAKDLA	MPGWK-----	HDAK	436				
RESULT 13									
AAW97386									
ID	AAW97386	standard; protein; 437 AA.							
XX	XX								
AC	AAW97386;								
XX	XX								
DT	14-MAY-1999	(first entry)							
XX	XX								
DE	A	fructosylamino acid oxidase.							
XX	XX								
KW	Fructosylamino acid oxidase; FAOD-P; Penicillium janthinellum S-3413.								
XX	XX								
OS	Synthetic.								
OS	Penicillium janthinellum.								
XX	XX								
FN	JP11046769-A.								
XX	XX								
PD	23-FEB-1999.								
XX	XX								
PP	05-AUG-1997; 97JP-00210609.								
XX	XX								
PR	05-AUG-1997; 97JP-00210609.								
XX	XX								
PA	(KYOT-) KYOTO DAIICHI KAGAKU KK.								
XX	XX								
DR	WPI; 1999-208112/18.								
DR	N-ESDB; AAX15949.								
XX	XX								
PT	New DNA coding fructosylamino acid oxidase - synthetically designed and recombinantly prepared.								
PT	Claim 1; Page 8-10; 2lpp; Japanese.								
PS	XX								
CC	The present sequence represents a fructosylamino acid oxidase (FAOD-P), and is encoded by a synthetic sequence. The wild type FAOD-P sequence								
CC	is derived from Penicillium janthinellum S-3413. The synthetic sequence								
CC	was constructed using primers AAX15951-6006								
XX	XX								
SQ	Sequence 437 AA;								
Query Match 26.1%; Score 660; DB 2; Length 437;									
Best Local Similarity 33.6%; Pred. No. 3.5e-56;									
Matches 156; Conservative 75; Mismatches 157; Indels 76; Gaps 14									
Qy	16	KSSLLIV-GAGTGTGTSTALHLARGYT--	NVTVLDPPVPSPAISAGNDVNKVISGGQS	72					
Db	6	ESTKIVIVGGGTWGSSTALHLIRSGYPSNITVLDVYPIPSLSAGYDLNKIMSI--	RL 63						
Qy	73	NKDEIEVNEILAEAFNGWKNDPLFKPYHYHDTLLMSACSGEGDRLGRVVRPGE	DNPL 132						
Db	64	RNGPDQQ---LSLEALDMWKNDPLFKPFPHNVGMDCSSSGEGIASLRKKHQDLIDANI	119						
Qy	133	-----VELTRPEQPRKLAPEGVLQDPPGKKGYFARSGAGWAHARNALVAAAREAQRMG	186						
Db	120	GLEKTNIIWESEDDILAKAPH-FTREQIKGWKGLFCGDG-GWLAAKAINAIGTFLKSGQ	177						
Qy	187	VKFVTGT-----PQGRVTVLIFENNDVKGAVTGDGKIWAERTFLCAGASAGQPL	236						
Db	178	VKFGGSGAGTFRPLFAPDGATCS-----GVETVDGTYKCFADKVVLAGANSS	TLV 228						
Qy	237	DFKNQLRPTAWTLVHIAKLPPEARLYKNIPIVFNTERGFFFPBPDSERGEIKICDHPGYT	296						

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:35:35 ; Search time 36.2303 Seconds
(without alignments)
972.511 Million cell updates/sec

Title: US-10-622-893A-5

Perfect score: 2530

Sequence: 1 MGSGDDDDALAVTKSSSL.....LEGLEPIENPLRTGHHHHH 472

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1982.5	78.4	437	2	US-09-031-059-1
2	1982.5	78.4	437	2	US-09-031-059-3
3	168	6.6	433	4	US-09-711-164-412
4	155	6.1	387	3	US-09-457-302-1
5	150	5.9	866	4	US-09-949-016-10219
6	146.5	5.8	448	4	US-09-252-991A-21043
7	136.5	5.4	383	4	US-09-893-096-5
8	135.5	5.4	378	3	US-09-199-637A-367
9	135.5	5.4	419	4	US-09-252-991A-21423
10	134.5	5.3	484	4	US-09-252-991A-24053
11	125	4.9	652	4	US-09-252-991A-30872
12	124.5	4.9	476	4	US-09-489-039A-12890
13	123.5	4.9	486	4	US-09-684-405-2
14	121.5	4.8	669	4	US-09-213-888-25
15	121.5	4.8	669	4	US-09-328-877D-25
16	117.5	4.6	374	4	US-09-468-647A-118
17	116.5	4.6	431	1	US-08-391-339-18
18	116.5	4.6	431	1	US-08-484-274A-18
19	115.5	4.6	568	4	US-09-266-965-135
20	115	4.5	505	4	US-09-902-540-9986
21	114.5	4.5	474	4	US-09-543-681A-7499
22	113.5	4.5	635	4	US-09-543-681A-7499
23	113.5	4.4	477	4	US-09-252-991A-27656
24	112.5	4.4	626	4	US-09-331-568A-27
25	110.5	4.4	375	4	US-09-328-352-6891
26	108	4.3	380	4	US-09-489-039A-7473
27	108	4.3	430	1	US-08-391-339-5

Sequence 5, Appli
Sequence 9, Appli
Sequence 7036, Ap
Sequence 81, Appl
Sequence 81, Appl
Sequence 14, Appl
Sequence 8, Appli
Sequence 2, Appli
Sequence 7529, Ap
Sequence 5943, Ap
Sequence 8, Appli
Sequence 18996, A
Sequence 32, Appl
Sequence 1, Appli
Sequence 26523, A
Sequence 12, Appl
Sequence 80, Appl
Sequence 20386, A

ALIGNMENTS

RESULT 1

US-09-031-059-1

; Sequence 1, Application US/09031059

; Patent No. 5948659

; GENERAL INFORMATION:

; APPLICANT: KATO, NOBUO

; APPLICANT: SAKAI, YASUYOSHI

; APPLICANT: TANI, YOSHIKI

; APPLICANT: FUKUYA, HIROSHI

; TITLE OF INVENTION: RECOMBINANT FRUCTOSYL AMINO ACID

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH

; STREET: PO BOX 747

; CITY: FALLS CHURCH

; STATE: VA

; COUNTRY: USA

; ZIP: 22040-0747

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/031,059

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/899,172

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: WEINER, MARC S

; REGISTRATION NUMBER: 32,181

; REFERENCE/DOCKET NUMBER: 0020-4253P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 205-8000

; TELEFAX: (703) 205-8050

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 437 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-09-031-059-1

Query Match 78.4%; Score 1982.5; DB 2; Length 437;
Best Local Similarity 82.2%; Pred. No. 3.4e-194;
Matches 360; Conservative 44; Mismatches 33; Indels 1; Gaps 1;

```
QY 12 LAVTKSSLLIIVGAGTGTSTALHLARRGYTNVTLDPPVPFSAISAGNDVNVKVISSGOY 71
DB 1 MPVTKSSLLIIGAGTGGCSTALHLARRGYTNVTLDPPVPFSAISAGNDVNVKIISSGOY 60
QY 72 SNKKBIEVNEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQEGLDRLGVRVPGEDPN 131
DB 61 SSKKBIEVNEILAEAFNGKNDPLFKPYHYDHTGVMGSAATQEGLERLGRVVRPEDEPD 120
QY 132 LVELTRPEQFKLAPGVLQDPPGKGYFARSGAGWAHARNALVAAAREAQRMGKVFY 191
DB 121 VAELTRPEQFROLAP-GVLKGNFPGWGRGYHRSNAGWAHARNALVAAAREAQRLGVRFA 179
QY 192 GTPQGRVVTILFENNNDVKAVTGDGKIWRAERTFLCAGASAGQFLDFKNQLRPTAWTLVH 251
DB 180 GSPQGRVITLIFENNNDVKAVTADGKIWRAEQITLCAGAAAGQFLDFKQQLRPTAWTLVH 239
QY 252 IALKPEERALKYNIPIVFNIERGFPEPDEERGEIKICDEHPGYTNMVSADGTMMISIPF 311
DB 240 IQLKPEERALKYNIPIVFNIERGFPEPDEERGEIKICDEHPGYTNMVSADGTMMISIPF 299
QY 312 EKTQIPKEAETRVALLKETMPQLADRPFSFARICWCADTANREFLIDRHPOVHSLVLCG 371
DB 300 EKTQVPEAREMVRKLLSETMPQLADRPFSFARICWCADTANREFIIDRHPEYPSLVLCG 359
QY 372 GASGRGFKYLPSTGNLIVDAMEGKVPQKIHELKWNPDIAANNRWDITLGRFGGPNRVM 431
DB 360 GASGRGFKYLPSTGSIADAMEDKTPAKIHKLRWSPEIARNWGDRLGRFGGPNRVM 419
QY 432 FHDVKWNTVOYRDISKL 449
DB 420 FNEVKWNTVQTDISKL 437
```

RESULT 2

```
US-09-031-059-3
; Sequence 3, Application US/09031059
; Patent No. 5948659
; GENERAL INFORMATION:
; APPLICANT: KATO, NOBUO
; APPLICANT: SAKAI, YASUYOSHI
; APPLICANT: TANI, YOSHIKI
; APPLICANT: FUKUYA, HIROSHI
; TITLE OF INVENTION: RECOMBINANT FRUCTOSYL AMINO ACID
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,059
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/899,172
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 0020-4253P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-059-3
Query Match 78.4%; Score 1982.5; DB 2; Length 437;
Best Local Similarity 82.2%; Pred. No. 3.4e-194; Indels 1; Gaps 1;
Matches 360; Conservative 44; Mismatches 33;
QY 12 LAVTKSSLLIIVGAGTGTSTALHLARRGYTNVTLDPPVPFSAISAGNDVNVKVISSGOY 71
DB 1 MPVTKSSLLIIGAGTGGCSTALHLARRGYTNVTLDPPVPFSAISAGNDVNVKIISSGOY 60
QY 72 SNKKBIEVNEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQEGLDRLGVRVPGEDPN 131
DB 61 SSKKBIEVNEILAEAFNGKNDPLFKPYHYDHTGVMGSAATQEGLERLGRVVRPEDEPD 120
QY 132 LVELTRPEQFKLAPGVLQDPPGKGYFARSGAGWAHARNALVAAAREAQRMGKVFY 191
DB 121 VAELTRPEQFROLAP-GVLKGNFPGWGRGYHRSNAGWAHARNALVAAAREAQRLGVRFA 179
QY 192 GTPQGRVVTILFENNNDVKAVTGDGKIWRAERTFLCAGASAGQFLDFKNQLRPTAWTLVH 251
DB 180 GSPQGRVITLIFENNNDVKAVTADGKIWRAEQITLCAGAAAGQFLDFKQQLRPTAWTLVH 239
QY 252 IALKPEERALKYNIPIVFNIERGFPEPDEERGEIKICDEHPGYTNMVSADGTMMISIPF 311
DB 240 IQLKPEERALKYNIPIVFNIERGFPEPDEERGEIKICDEHPGYTNMVSADGTMMISIPF 299
QY 312 EKTQIPKEAETRVALLKETMPQLADRPFSFARICWCADTANREFLIDRHPOVHSLVLCG 371
DB 300 EKTQVPEAREMVRKLLSETMPQLADRPFSFARICWCADTANREFIIDRHPEYPSLVLCG 359
QY 372 GASGRGFKYLPSTGNLIVDAMEGKVPQKIHELKWNPDIAANNRWDITLGRFGGPNRVM 431
DB 360 GASGRGFKYLPSTGSIADAMEDKTPAKIHKLRWSPEIARNWGDRLGRFGGPNRVM 419
QY 432 FHDVKWNTVOYRDISKL 449
DB 420 FNEVKWNTVQTDISKL 437
```

RESULT 3

```
US-09-711-164-412
; Sequence 412, Application US/09711164
; Patent No. 6589738
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THEREOF
; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711,164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 412
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-711-164-412
Query Match 6.6%; Score 168; DB 4; Length 433;
Best Local Similarity 23.0%; Pred. No. 3e-08;
Matches 120; Conservative 59; Mismatches 176; Indels 166; Gaps 28;
QY 1 MGGSGDDDDDLALAVTKSSLLIIVGAGTGTSTALHLARRGYTNVTLDPPVPFSAISAGN 60
DB 7 LGHDMEDD-----CDIIIGAGTAGTACALCARAG-LSVLLLEAEIFG----- 50
QY 61 DVNKVSSGQYSNNKKBIEVNEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQEGLDRL 120
```

Db 51 --SKULSGRL-----YTHALAE-----LLQFHLTAPERRITHESL 89
Qy 121 ---GVRVRPGEENLVELTRPEQKRLAPEG-----VLOGDFPGWKGYFARSAGAHARN 173
Db 90 TPDGVTT-----PSSLQPGGESVLRARFDW----- 117
Qy 174 ALVAAREAQRMGVKFTGTGPOGRVVTLLPENNDKAGVTDGDKIWAERTFLCAGASAG 233
Db 118 -LVA---EABKEGVECI---PGATVDALYENGVRGCVICGD-DILRARVYVLAEGA--- 166
Qy 234 QFLDFKNQLRPTAWTLVHIALKPERALYKNIPIVFN-----IERGFFPEDEERGEI-- 286
Db 167 -----NSVLAERHGLV---TRPAGEAMALGIKEVLSLETSABERHLENNEGAALLFS 217
Qy 287 -KICDEHPG---YTNMVQADG-----TMMISIPFEKTQIPKEAETVRALLKET-- 331
Db 218 GRICDDLPGCAFLYTNQTLGLGVCPLSSLTQSRVPASELLTRFKAPAVRPLIKNTES 277
Qy 332 -----MPQ--LADRFSPARICW-----CAETANREFLIDRHPQVHSLVLCGA 373
Db 278 LEYGAHLVPEGLHSMFPVAGNGMLLVGDALRSCVNT-----GI 317
Qy 374 SGRGFKY---LPSIGNLIVDAMEGKVPQKIHELKWNPDIAANRNWRDLGRFGG-PNR 428
Db 318 SVRGMDWALTGAQAAQTLSACQHPQLFPL--YHNVERSLLW-DVLQRYQHVPL 374
Qy 429 VMDPHDKVNTVOYRDISKLKGELEGLPIPNPLLRTHHH 469
Db 375 LQRPGRVTRTPALM-QDISRLMDQGDKVP-PLRQLFWHH 413

RESULT 4

US-09-457-302-1
; Sequence 1, Application US/09457302
; Patent No. 6228626
; GENERAL INFORMATION:
; APPLICANT: ICHIKAWA, Toshio
; APPLICANT: KOYAMA, Yasuji
; TITLE OF INVENTION: SARCOSE OXIDASE AND PROCESS FOR PRODUCING THE SAME
; FILE REFERENCE: 081356/0139
; CURRENT APPLICATION NUMBER: US/09/457,302
; CURRENT FILING DATE: 1999-12-09
; EARLIER APPLICATION NUMBER: JP 10-354482
; EARLIER FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-457-302-1

Query Match 6.18; Score 155; DB 3; Length 387;

Best Local Similarity 22.8%; Pred. No. 5.4e-07; Indels 78; Gaps 22;
Matches 94; Conservative 62; Mismatches 178;

Qy 20 LLIVGAGTGTSTALHLAGRGYTNVTLDPVPVPSAISAGNVNKNVSSGOYSNNKDEIE 79
Db 7 VIVVGAGSMGAAGYILAKQGVKTLV-DSFDPHTNGSHGDTIRIHA-YGEGREYVP 64
Qy 80 VNEILAEAFNGWKNDFPKPYHDTGLM-----SACSQEGLDRGLGVRVRGEDPNLV 133
Db 65 F-ALRAQELVYELEKETHHK-IFTQTGVLVYGPKGSAFVSETM-----EAAIIH 112
Qy 134 ELTRPEQKRLAPEGVLQDFPGWK-----YFA--RSGAGAHARNALVAAAREQRM 185
Db 113 SLEH-ELF-----EGKQLTD--RWAGVEVPDNYEAFIEPNSGVLFSENCIOAYRELAEH 164
Qy 186 GVRFTVGTGP--QGRVVTLLPENNDKAGVTDGDKIWAERTFLCAGASAGQFLDFKNQLR 243
Db 165 GATVLTTPVEDFEVTEDLVITAKGYSY-----ANKLVSMGAWNSKLSKLD--- 214

Qy 244 PTAWTLVHIALKPE-----ERALYKN---IPV-IFNIERGFPPE-PDEERGEIKIC 289
Db 215 -----VEIPLQYRQVVGFECEDEAKYSNNAHPAFMVEVENGIYGFPSFGSGGLKI- 267
Qy 290 DEHPGYTNMQSADGTMMISIPFEKTQIPKEAETVRALLKETMPQLADRPSPFARICWCA 349
Db 268 ----GYHSYGOQIDPDITINREFGAY---PEDEANLRKFLBQYMPG-ANGELKKGAVCMYT 319
Qy 350 DTANREFLIDRHPQVHSLVLCGASGRGFKYKPLPSIGNLIVD-AMEGKVPQKI 400
Db 320 KTDPEHEFVIDLHPKYSNVAIAAGFSGHGFKFSVVGETLAQLATTGTKEHDI 371
RESULT 5
US-09-949-016-10219
; Sequence 10219, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20/237,768
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10219
; LENGTH: 866
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10219

Query Match 5.9%; Score 150; DB 4; Length 866;

Best Local Similarity 20.8%; Pred. No. 6.4e-06;
Matches 102; Conservative 77; Mismatches 187; Indels 124; Gaps 23;

Qy 2 GSGDDDDDLALAVTK---SSLLIVGAGTGTSTALHLAGRGYTNVTLDPVPVPSAISA 58
Db 31 GREGEKPPLSAETQMKDRAETVIIGGCGVSLAYHLAKAGMKDVVLE----KSELTA 86
Qy 59 GN-----DVNKVISSGOYSNNKDEIEVNEILAEAFN-GWKNDFLPKPY 101
Db 87 GSTWHAAGLTYTHPGINLAKI-----HYDSIKLYELEETGVQVG----- 128
Qy 102 YHDTGLMSACSQEGLDRGLRV-RPGEDPNLVELTRPEQKRLAP----EGVLQGDPPG 156
Db 129 FHQPGSIRLATTPVRVDEFKYQMTRTGWHATEQYLIETPKIEMFPLNNMKNVLAGLY-- 186
Qy 157 WKGYPARSAGAHARNALVAAAREAQRMGVKFTGTGPOGRVVTLLPENNDKAVTGDG 216
Db 187 -----NPGDGHITDPVSLTWALAAGARKCAGALLKYPAP---VTLSKARSOGTVDVETPQG 237
Qy 217 KIWAERTFLCAG---ASAGQFLDFKNQLRPTAWTLVHIALKPERALYKNIPIVFNIE- 272
Db 238 SM-RANRIVNAAGFWAREVGNMIGLEHPLIPVHQYVVTSTIPEVKALKRELPLVRLLEG 296
Qy 273 -----RGFFPEDEERGEIKICDE-----HPGYTNMV--QSADGTMMISIPFEKTQI 316
Db 297 SYVLRQERDGLLFGPYESQEKMKVQDSWVTNGVPPGFGKELFESDLDRIMEHI----- 349
Qy 317 PKAETVRALLKETMPQLADRPSPFARICWCADTANREFLIDRHPQVHSLVLCGASGR 376
Db 350 -KAAMENVPVLKKADIINVVNGPITYS-----PDILP---MVGPHQGVRYNVAIG-FGY 399
Qy 377 GFKYLPSIGNLIVD-AMEGKVPQKIHELKWNPDIAANRNWRDLGRFGGPNRVMDPHDV 435
Db 400 GIHAGGVGKYLSDNLIHGEPP---FDLIEDPN-----RYG----- 433

```

QY 436 KWTNVQYRD 445
      :|||:
Db 434 -KWTTOYTE 442

RESULT 6
US-09-252-991A-21043
; Sequence 21043, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21043
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21043

Query Match 5.8%; Score 146.5; DB 4; Length 448;
Best Local Similarity 23.7%; Pred. No. 5.1e-06;
Matches 98; Conservative 59; Mismatches 164; Indels 93; Gaps 23;

QY 18 SLLIVGAGTGTSTALHLARRGYTNVTLDPPVPSAISAGNDVNKVISSGOYSNNKDE 77
      : : : ||| ||| ||| : ||| : ||| : : : : ||| :
Db 51 TDCVIGAGYTGSLTALFLENGF-KVTVLETAKVGFGAS-GRNGGOIVNS--YSRDIDV 106

QY 78 IEVN-----EIIAEEAFNGMNDPLPKPYHDTGL-----LMSACSQEGDLRLGVR 123
      : : : : : : : : : : : : : : : : : : : : : : :
Db 107 IERTVGKRAQLLGEMAFEGR---IIRVARVGIQCDLKDGGVFAAFTEKQMDHLRAQ 163

QY 124 VRPGE--DPNLVELTRPEQPKLAPEGVLODPPGWKGYFARGAGWAHARNALVAAARE 181
      : : : : : : : : : : : : : : : : : : : : : : :
Db 164 KQLWERYGHQLEITMDAKRIR----EYVATDNYIG--GMLDMSG-GHIHPLNALGEAAA 216

QY 182 AORGVGVFTGTPOGRVVTLIFENNVDKGAVTGDKIWRABERTFLCAGASAGOFDFPNQ 241
      : : : : : : : : : : : : : : : : : : : : : : :
Db 217 VESLGGRIRYQSPATRI-----ERGASPVVHTPGKV---KAFIIVVAGNA--YLN---- 262

QY 242 LRPTAWTLVHIALKPEERALKYKNIP-----VFNIERGFFPEPDEERGEIKICDEHP 293
      : : : : : : : : : : : : : : : : : : : : : : :
Db 263 -----GLVPPELAA--KSMPCGQVITTEPLNEELAHSLLPQD-----YCVEDC 303

QY 294 GYTNWVOSADGTWMSI-----PFKEIQIPKEAETFRALLKETTWPOLADRPFSFARICWCA 349
      : : : : : : : : : : : : : : : : : : : : : : :
Db 304 NYLLDYRLSGDKRLIYGGGVYIGARDPADIEATIRPKMLKTPQLKDVKIDFA---W-- 358

QY 350 DTANREFLIDRHPO-----YHSLVLGGCAGRGPKYLPSPIGNLITVDAMEGK 395
      : : : : : : : : : : : : : : : : : : : : : : :
Db 359 -TGNFLTSLSLFPQVGRIGDNIYISQ-----GCSGHGVYTHVAGKVLAEALRGQ 407

RESULT 7
US-09-883-096-5
; Sequence 5, Application US/09883096
; Patent No. 6680369
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Craven, Andrew
; APPLICANT: Yu, Ming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE

```

```
; SEQ ID NO 367
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-199-637A-367

Query Match      5.4%; Score 135.5; DB 3; Length 378;
Best Local Similarity 24.3%; Pred. No. 5.1e-05;
Matches 102; Conservative 45; Mismatches 165; Indels 107; Gaps 19;

Qy 17 SSSLLIVGAGTGTSTALHARRGTYNTVLDPPVPSAISAGNDVNVKVISSGQYNNKD 76
Dy 16 SRDVVVVGAGVIGLLTARELALAGLRVTLVERGESGREASWAGGIVSPLYPRYSP--- 72
Qy 77 EIEVNEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQEGLDRIGVRVPCEDNPL----- 132
Dy 73 -----AVTALAHWSQD---FYP-----ALQORLLDETGL-----DPEVHTVG 106
Qy 133 ---VEL-----TRPEQFRKLAPGEGVLQDPPGKWKGYFARS--GAGWAHARNA 174
Dy 107 LYWLDDDDQTEALQWARKHTRP---LKEVPFEEAYAAVPGLGAGFQRAVYMSGVANVRNP 163
Qy 175 LVAAAREAQRMGVKFTGTGQGRVVVTILFENNVDKGAVTGDKIWRABERTFLCAGASAGQ 234
Dy 164 RLARSLRASLQOQFANLELHEQTEVRGWLDRDGVGVATSRGEI-RGDKVLLAAGAWSGE 222
Qy 235 FLDFKQNLQRPATWTLVHIALKPEERALKYK-----NIPVFNIERGFFFPDDEBERGEIKCD 290
Dy 223 L-----LKPGLGLELPVVPVK-QOMILYKCAADFLPRMV-LAKGRYAIPRRD-GHILIGS 273
Qy 291 --BHPGYTNMVQSADGTMTMSIPPEKTOIPEAEATRVRLALKETMPOLAD-----R 338
Dy 274 TLEHSG-----FDKTPDDEALES-LRASAAELLPELADMQPVNAHWAGLR 316
Qy 339 PFSPARICWCADTANREFLIDRHPQYHSLVLCGCSGRGFKYLPISIGNLIIVDAMEGKVP 397
Dy 317 PGSPEGIPY-----ICPVPGDGLWNTGHYRNLGLVLAFLAPASCELLADLMSGREP 365

RESULT 9
US-09-252-991A-21423
; Sequence 21423, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21423
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-21423

Query Match      5.4%; Score 135.5; DB 4; Length 419;
Best Local Similarity 24.3%; Pred. No. 6.1e-05;
Matches 102; Conservative 45; Mismatches 165; Indels 107; Gaps 19;

Qy 17 SSSLLIVGAGTGTSTALHARRGTYNTVLDPPVPSAISAGNDVNVKVISSGQYNNKD 76
Dy 57 SERVVGAGVIGLLTARELALAGLRVTLVERGESGREASWAGGIVSPLYPRYSP--- 113
Qy 77 EIEVNEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQEGLDRIGVRVPCEDNPL----- 132
Dy 114 -----AVTALAHWSQD---FYP-----ALQORLLDETGL-----DPEVHTVG 147
```

```
Qy 133 ---VEL-----TRPEQFRKLAPGEGVLQDPPGKWKGYFARS--GAGWAHARNA 174
Dy 148 LYWLDDDDQTEALQWARKHTRP---LKEVPFEEAYAAVPGLGAGFQRAVYMSGVANVRNP 204
Qy 175 LVAAAREAQRMGVKFTGTGQGRVVVTILFENNVDKGAVTGDKIWRABERTFLCAGASAGQ 234
Dy 205 RLARSLRASLQOQFANLELHEQTEVRGWLDRDGVGVATSRGEI-RGDKVLLAAGAWSGE 263
Qy 235 FLDFKQNLQRPATWTLVHIALKPEERALKYK-----NIPVFNIERGFFFPDDEBERGEIKCD 290
Dy 264 L-----LKPGLGLELPVVPVK-QOMILYKCAADFLPRMV-LAKGRYAIPRRD-GHILIGS 314
Qy 291 --BHPGYTNMVQSADGTMTMSIPPEKTOIPEAEATRVRLALKETMPOLAD-----R 338
Dy 315 TLEHSG-----FDKTPDDEALES-LRASAAELLPELADMQPVNAHWAGLR 357
Qy 339 PFSPARICWCADTANREFLIDRHPQYHSLVLCGCSGRGFKYLPISIGNLIIVDAMEGKVP 397
Dy 358 PGSPEGIPY-----ICPVPGDGLWNTGHYRNLGLVLAFLAPASCELLADLMSGREP 406

RESULT 10
US-09-252-991A-24053
; Sequence 24053, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24053
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-24053

Query Match      5.3%; Score 134.5; DB 4; Length 484;
Best Local Similarity 22.4%; Pred. No. 9.7e-05;
Matches 103; Conservative 70; Mismatches 185; Indels 101; Gaps 23;

Qy 12 LAVTKSSLLIVGAGTGTSTALHARRGTYNTVLDPPVPSAIS-----AGND 61
Dy 77 LAGEEQADVCVGGGFGVNTALELAQGRF-SVLLLEAHRIGMGASGRNGQLIRGVGHD 135
Qy 62 VNK---VISSGQYNNKNDIEVNEILAEAFNGKNDPLFKPYHYDHTGLLMSACSQEGLD 118
Dy 136 VEQFFPVIGA-----DGVKALKMLGLEAVEIVR--RVEQYVADCDLKWGYCD----- 181
Qy 119 RLGRVVRPG-----EDPNLVELTRPEQFRKLAP-----EGVLQDPPGKWKGYFARSAGW 168
Dy 182 ---LANKPGDYQGFREDMEELQALGYRHEMLVPAEEMRSVVGSD--RYVGGGLVDMGSGH 236
Qy 169 AHARNALVAAAREAQRMGVKFTGTGQGRVVVTILFENNVDKGAVTGDKIWRABERTFLCA 228
Dy 237 LHPNLNLVLGAAAAQSLGVRLEPERSP---VTRIDYGAEQV-VHTATGKV-RAKT--LVL 288
Qy 229 GASAGQFLDFKQNLQRP-----TAWTLVHIALKP-EERALYKNIPVFNIERGFFFPD 280
Dy 289 GCNA-----YMNDLNPLLGKVLPAQSYV-IAEPLDEQLARQLLP----- 328
Qy 281 EERGEIKICDEHPGYTNMVQSADGTW--SIPPEKTQIPKEAETRVRLALKETMPOLADR 338
Dy 329 ---QNMVAVCDQORVALDYRLSADNRLFGGACHYSGRDPSPDIAAYMRPKMLEVFPQLANV 385
Qy 339 PFSPARICWCADTANREFLIDRHPQYHSLVLCGCSGRGFKYLPISIGNLIIVDAMEG--- 394
```

```
Db 386 RIDYQGMIGIGANRLPQIRPGQPNVYFAQYSGHGNATHLAGQLLAEAGGQSD 445
QY 395 -----KVPOKIR---ELIKWNPDIANRNW---RDTLG 421
Db 446 GFDLFAKVPHTPGGKLLR-SPLLALGMAWYRLKEKLG 483

RESULT 11
US-09-252-991A-30872
; Sequence 30872, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30872
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30872

Query Match 4.9%; Score 125; DB 4; Length 652;
Best Local Similarity 21.8%; Pred. No. 0.0015;
Matches 92; Conservative 54; Mismatches 162; Indels 114; Gaps 16;

QY 17 SSSLLVAGCTGTSTALHARRGYNVTVDLP-YPVPSAISAGNDVNVKVISSGQYNNK 75
Db 283 NAEVIVGAGIVGSACAHELARRG-LDVLVDSRRGGATAVGMGLV-----AMDDNP 334
QY 76 DETEVNEILAEAFNGWKNPDKPKPYVHDGGLMSAC-----SOGELDRLGVRVR 125
Db 335 AELASD-YSIQAWRTWAADLPEDCAYRNGCTLWLADAELAEAEKRAALLAGV--- 390
QY 126 PGEDPNLVELTRPEQFKLAPEGVLQDFPGWKGYFARSGAGWAHARNALVAAAREAQRM 185
Db 391 -----ACEMLDAARLDLEP-VLR---PGLAGALKVPDGDILYAPN---AARWLLERA 436
QY 186 GVKFVTGTPGQVVTLLFENNVDKGA--VTGDGKIWAERTFLCAGASAGQFLDFKNQLR 243
Db 437 G-----PRLRLLHAEVSEVDGSRRLADGRWLSAEALVLANGIHAGEL----- 479
QY 244 PTAWTLVHIALKPEERALKYKNIPIVFIENRGGFFPEDEERGEIKICDEHP----- 293
Db 480 -----CAELPIRP-----KGLHLLITDYPGTLRHQLVEL 509
QY 294 GYTNMVQASDGTWM-----SIPFEKTQIPKEAETRVALLK--ETWQOL 335
Db 510 GYVSSAHASSGTSVAFNAQPRPTGQVFLGSSRQFDLDPQVEGVLARMLRRALLDYLPL 569
QY 336 ADPRFPFARICWADTANREFLDHQPYHSLVLCGASGRGFKYLPISGNLIVDAMEGK 395
Db 570 AGLNAIRAWTGFRAATPDGLPLGHEHPAQPGFLWLVAGHEGLGVTTPAGSARLLAAQLFGE 629
QY 396 VP 397
Db 630 TP 631

RESULT 12
US-09-489-039A-12890
; Sequence 12890, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
```

```
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12890
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12890

Query Match 4.9%; Score 124.5; DB 4; Length 476;
Best Local Similarity 22.6%; Pred. No. 0.001;
Matches 94; Conservative 48; Mismatches 153; Indels 121; Gaps 21;

QY 11 ALAVTKSSLLVAGCTGTSTALHARRGYNVTVDLPY-----PVPISAISA 58
Db 66 ALREDIADVVIITGGFGSINTALEAEQITNVVLEARHLGYGCTGRNGGQVMAGI-- 123
QY 59 GNDVNVKVISSGQYNNKDEIE-----VNEILAEAFNGWKNPDKPKPYVHDGGLMSA 111
Db 124 GHDIKAV---KXHVKGEGLETLFKIANLGAGIIRIRKYNIDADFVPGY---GYL--A 174
QY 112 CSOEGELDRLG-----VRVRPGEDPNLVELTRPEQFKLAPEGVLQDFPGWKGYFARSG 165
Db 175 YNQRQLKTLRQWEKEFKATPDDE---IELYTGKEVQQVVGSEV-----YCGALKHMG 224
QY 166 AGWAHARNALVAAAREAQRMGVKFTVGTGTP-----QGRVVTLLFENNVDKGAVTGDGK 217
Db 225 GGQIHSNMLLSGAQAASHLGVKIPFESSPVVEVNYGKQVRVRTAM-----GSVKAACL 277
QY 218 IWRPAAETFLCAGASAGQFLDFKNQLRPTAW--TLV-----HIALKPEERALKYKNIPIVFI 271
Db 278 LWACD-----SFLNNMEPEIYNKTLVTYSYQVSTEPSDELIERISPL--- 320
QY 272 ERGFFPEDEERGEIKICDEHPGYTNVQASADGTMTMSIPFEKTQIPKEAETRVALLKET 331
Db 321 -RGAFSDIRPVINYVYRTRE---NRLJFGSATRF---VEYT--PNDFAWNRTLLAEV 369
QY 332 MPOLADRPFSFARICWADTANREFLDHQPYHSLVLCGASGRGFKYLPISGNL 387
Db 370 FPYLRDVKIDFA---NGGPMA-----CSAN-----LFPQIGTL 399

RESULT 13
US-09-684-405-2
; Sequence 2, Application US/09684405
; Patent No. 6770477
; GENERAL INFORMATION:
; APPLICANT: Dennis J. Slamon
; APPLICANT: Juliana J. Oh
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED GENES
; FILE REFERENCE: 30448.79USU1
; CURRENT APPLICATION NUMBER: US/09/684,405
; CURRENT FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/157,923
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-684-405-2

Query Match 4.9%; Score 123.5; DB 4; Length 486;
Best Local Similarity 20.9%; Pred. No. 0.0013;
Matches 94; Conservative 53; Mismatches 179; Indels 123; Gaps 20;

QY 18 SSSLVAGCTGTGTSTALH-----ARRGYNVTVDLPYVPISAISAGNDVNVK----- 65
```

```
Db 64 SDVVVGGVGLSVATWKLKLESGRAIRVLVVERDHTYSQASTGLSVGGICQOFSLPE 123
Qy 66 -----ISSGQYNNKNDIEVNEILAEBAFNKNDPLFKPYHYHTDGLMSACSGEGLD-R 119
Db 124 NIQLSLSFASFLRN-----INEYLAV-----VDAPPLDLRFNPSGYLLLASAKDAAME 172
Qy 120 LGVVRGEGDNLVELTPEQFRKLAP-----EGVLQDFFGKGYFARSAGNAHARNAL 175
Db 173 SNKVQVQEGAK-VSLMSPDQLRNKFPWINTGVALAS-----YGMEDGWFDPWCLL 224
Qy 176 VAAAREARMGVKEVTG-----TPQGRVVTL-----IPENNVDVG 210
Db 225 QGLRRKVQSLGVLEFCQEVTRFVSSQRMILTDDKAVLKRHEVHVMDRSLEYQVPEC 284
Qy 211 A-VTGDKINRAERTFLCAGASAGQFLDFKNQLRPTAWTLVHIALKPEERALKYKNIPVIF 269
Db 285 AIVINAAGAWSAQIAAL-AGVGEGP-----PGTLQGTKLPEPRKRVY----- 327
Qy 270 NIERGPFEPDDEBERGEIKICDEHPGY-TNMVQSADGTW-----MSIPFEKTOIPKEATR 323
Db 328 -----VMHCPQGGLETPLVADTSGAYFRREGLSNYLGGRSPTEQEEP 371
Qy 324 VRALLK-----ETWPLADRPFSFARI-----CWCA-----DTANREFLIDRHPQVHSL 367
Db 372 DPNANLEVDHDFQDKVWPHLALRVAPETIKVQSAWAGYDYNTFDQGVVGHPLVVM 431
Qy 368 VLGGASGRGFKYLPSTGNLIVD-AMEGK 395
Db 432 YFATGFSGHGLQQAAPGIGRAVENVLKGR 460
```

RESULT 14

```
US-09-213-888-25
; Sequence 25, Application US/09213888A
; Patent No. 6638731
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213.888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: V5HIS tagged
; OTHER INFORMATION: homo sapien
US-09-213-888-25
```

```
Query Match 4.8%; Score 121.5; DB 4; Length 669;
Best Local Similarity 51.9%; Pred. No. 0.0035;
Matches 27; Conservative 6; Mismatches 10; Indels 9; Gaps 4;
```

```
Qy 429 VMDPFH-DVKEWTN--VOYRDISKLKGEGLPIPNPLL-----RTGHHHHHH 472
Db 619 VLDVDVDMKEFCRYPAQWRPLES-RGPFEGKPIPNPLLGLDSTRTGHHHHHH 669
```

RESULT 15

```
US-09-328-877D-25
; Sequence 25, Application US/09328877D
; Patent No. 6730778
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
```

```
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328.877D
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: V5HIS tagged
; OTHER INFORMATION: homo sapien
US-09-328-877D-25
```

```
Query Match 4.8%; Score 121.5; DB 4; Length 669;
Best Local Similarity 51.9%; Pred. No. 0.0035;
Matches 27; Conservative 6; Mismatches 10; Indels 9; Gaps 4;
```

```
Qy 429 VMDPFH-DVKEWTN--VOYRDISKLKGEGLPIPNPLL-----RTGHHHHHH 472
Db 619 VLDVDVDMKEFCRYPAQWRPLES-RGPFEGKPIPNPLLGLDSTRTGHHHHHH 669
```

```
Search completed: May 27, 2005, 13:50:01
Job time : 37.2303 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:34:21 ; Search time 36.2303 Seconds
(without alignments)
1253.491 Million cell updates/sec

Title: US-10-622-893A-5
Perfect score: 2530
Sequence: 1 MGSSGDDDLALAVTKSSSL.....LEGLPINPLRTGHHHHH 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	488.5	19.3	412	2 T40295	fructosyl amine -
2	465	18.4	433	2 T37605	probable fructosyl
3	192	7.6	390	2 I39975	sarcosine oxidase
4	181	7.2	389	2 JS0671	sarcosine oxidase
5	177.5	7.0	390	2 JC7256	L-pipecolate oxida
6	170	6.7	389	2 T44248	sarcosine oxidase
7	168	6.6	433	1 B65058	fixC protein homol
8	155	6.1	387	2 JU0461	sarcosine oxidase
9	151.5	6.0	433	2 E91081	hypothetical prote
10	151.5	6.0	433	2 F85926	hypothetical prote
11	148	5.8	380	2 AB3160	santhopine deamina
12	146.5	5.8	427	2 G83298	conserved hypothet
13	142.5	5.6	857	2 S16133	dimethylglycine de
14	139.5	5.5	364	2 B83078	probable D-amino a
15	138	5.5	376	2 A83905	hypothetical prote
16	137.5	5.4	372	2 A85668	sarcosine oxidase-
17	137.5	5.4	372	2 E90808	sarcosine oxidase-
18	136.5	5.4	372	2 JC5371	probable sarcosine
19	136.5	5.4	382	2 G75122	sarcosine oxidase,
20	136.5	5.4	416	2 AD3075	sarcosine oxidase
21	136.5	5.4	416	2 E98211	sarcosine oxidase
22	136.5	5.4	825	2 F95963	probable dehydroge
23	135.5	5.4	372	2 F90159	sarcosine oxidase,
24	132	5.2	371	2 F83487	hypothetical prote
25	129.5	5.1	439	2 E82983	probable oxidoredu
26	128.5	5.1	382	2 F71008	probable sarcosine
27	127.5	5.0	433	2 A10882	probable oxidoredu
28	127	5.0	372	2 A10637	probable sarcosine
29	127	5.0	372	2 A70019	opine catabolism h

30	126	5.0	652	2 AH2245	thiamin biosynthes
31	125.5	5.0	417	2 G98200	sarcosine oxidase
32	125.5	5.0	417	2 AB3086	sarcosine oxidase
33	125.5	5.0	442	2 D97382	hypothetical prote
34	125.5	5.0	442	2 AC2600	oxidoreductase ord
35	123.5	4.9	437	2 T45533	agaE protein (impo
36	116.5	4.6	452	2 G95356	probable oxidoredu
37	114	4.5	396	2 AF2835	sarcosine oxidase
38	114	4.5	413	2 A97613	sarcosine oxidase
39	113	4.5	420	2 AC3284	D-amino-acid dehyd
40	113	4.5	703	2 T15503	hypothetical prote
41	112.5	4.4	629	1 BVECOA	glucose-inhibited
42	112.5	4.4	629	2 C91214	glucose-inhibited
43	112.5	4.4	629	2 D86060	glucose-inhibited
44	112	4.4	405	2 I40887	sarcosine oxidase
45	111.5	4.4	441	2 F98306	agaE protein (impo

ALIGNMENTS

RESULT 1

T40295
fructosyl amine - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T40295
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, March 1998
A;Reference number: Z21918
A;Accession: T40295
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-412 <WOO>
A;Cross-references: UNIPROT:O43029; EMBL:AL022071; PIDN:CAAL7815.1; GSPDB:GN00067; SPDB:
A;Experimental source: strain 972h-; cosmid c354
C;Genetics:
A;Gene: SPDB:SPBC354.15
A;Map position: 2

Query Match		19.3%;	Score 488.5;	DB 2;	Length 412;
Best Local Similarity		30.1%;	Pred. No. 1.3e-30;		
Matches 130;		Conservative 78;	Mismatches 183;	Indels 41;	Gaps 12;
Qy	14	VTKSSLLIYGAGTGWGTSTALHARRGYTNVTLDYPVPSAISAGNDVNKVISSGQYSN	73		
Db	1	MVKNTSVIIIVGAGVGLSAALETKRGYTIKILDRAPPVVDGSSVDANRIIRS-DYAD	59		
Qy	74	NKDEIEVNEILAEAFNGKNDPLFKPYHYDTGLLSA-----CSOEGDLRLGYRVR	125		
Db	60	-----AVYCSMGIDALEEENRNLFLKEQYFGSLMFVGRDNVEYRDMLENLTGMV---	111		
Qy	126	PGBDPNLVELTRPEQPKLAPEGVLQDFFPGWKGYFARSGAGWAHARNALVAAAREAQRM	185		
Db	112	-----SAKQOTTEELRLKLPKWI--GELNDGEAGYANFSSGWNARQSVKVVNYLAHA	164		
Qy	186	GVKFTVTGPGRVVTVLIFENNVDKGAVTGDKIWRABRTFLCAGASAGQFL-DFKNQLRP	244		
Db	165	GVSFISG-PEGTVEELITEENVVKGVRTTG-AVMAEKLIFATGAWTASLLPNDRFLA	222		
Qy	245	TAMTLVHIALKPERALYKNIPVIFNIERGFPEPDEERGEIKICDEHPGTYNNVQSADG	304		
Db	223	TGQPVAVIKLTPEYIRFLTNPNVYLDFTGTFYIPPTPDGYLKFARHGYGFTRMQLKSG	282		
Qy	305	TWMSIPEKTOI---PKEATRVALLKRTM-POLADRPSPFARICWCADTANREFLID	359		
Db	283	KVSVPPKKDLVPILPKAEALDRLRLQRTYGGEEISQRPFYKTRI CYTTDAAEFVFD	342		
Qy	360	RHPQYHSLVLCGASGRGFKYLPISIGNLIVDAMEGKVPQKIHILIKW---NPDIAAARNW	416		
Db	343	YHPDYENLFVCTGSGGHGFFPFLIGKYSIGCMFRELEELPKKRWKKNLEFAALDHS	402		
Qy	417	RDTLGRFGGPNR	428		

Db 403 R-----AGFSR 408 ||:|

RESULT 2

T37605

probable fructosyl amino acid oxidase - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T37605

R;McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Saunders, D.; Harris, D. submitted to the EMBL Data Library, October 1999

A;Reference number: Z21729

A;Accession: T37605

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-433 <CD>

A;Cross-references: UNIPROT:Q9UTW9; EMBL:AL132667; PIDN:CAB59618.1; GSPDB:GN00066; SPDB:

A;Experimental source: strain 972H-; cosmid c139

C;Genetics:

A;Gene: SPDB:SPAC139.04C

A;Map position: 1

Query Match 18.4%; Score 465; DB 2; Length 433;

Best Local Similarity 30.1%; Pred. No. 1e-28;

Matches 130; Conservative 80; Mismatches 174; Indels 48; Gaps 13;

QY 17 SSELLVAGTGTSTALHILAR-RGYTNVTLDPYVPSPSAISAGNDVNKVISSGOYNNK 75

DB 2 SRTIVVGGCVGLSTAVELAKNHSFDNIIAIDAEPPVSSMAANDINKIVRP-EYADLK 60

QY 76 DEIEVNEILAEAFNGWKNDFPKPYVHTDTGLMSACSQEGLDRLGVRVRPGE----- 128

DB 61 -----YMKLALEAMEKWRNDPELSSVYFEGRLST-----ISKDYPRAEFDEVAQRNL 109

QY 129 -----DNLVLTPEQPRKLAP-----EGVLQGFPGWKGYFARSGAGWAHARNALVAAR 180

DB 110 KLLGDSALINLSSEIRKKYPSLFSNSPLRSDMQA-----VNEHAGYANSAASLKLEL 165

QY 181 EAQRMGVKFTYGTPOGR-----VV-----TLIFENDVKGAVTGDKIWRARTFLCAGASA 232

DB 166 KARELGVEFFVG-KAGFKFKFVNNHSETDIDKDNHVSQVTEGTIYHADITILLAVAYL 224

QY 233 GQFLDFKNQRLPTAWTLVHIALKPEERALKNPVFNIERGFFFPDEBERGEIKICDEH 292

DB 225 NAYLNTSHRVCAGLPAHQLTDEEPTKYKNWPIIFDPCAVAFPPYPTKLIKLASTG 284

QY 293 PGYTNMVQS---ADGTTMSTIPF---EKTQIPKAEATRVALLKETMPQLADRPSPFARIC 346

DB 285 YEYVCNVETDYDENSKVSIPIHSGPSKSLPKYALIQMRRLDTFLPDLADRSILINTKMC 344

QY 347 WCADTANREFLIDHPQYHSLVLCGASGRGFKYLPISGNLIVDAMEGKVPQKIHILIKW 406

DB 345 WISDTEDANFLIDKVPQDFNVFANGDSGHAFKFLPNIGRIYAQRILGLDISEWKAARW 404

QY 407 NPDIANR-NWR 417

DB 405 REDDKASELKWR 416

RESULT 3

I39975

sarcosine oxidase (EC 1.5.3.1) - Bacillus sp.

C;Species: Bacillus sp.

C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004

C;Accession: I39975

R;Suzuki, K.; Imamura, S.; Sugiyama, M.

J. Ferment. Bioeng. 77, 231-234, 1994

A;Title: Cloning, sequencing, overexpression in Escherihia coil of a sarcosine oxidase-e

A;Reference number: I39975

A;Accession: I39975

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-390 <RES>

A;Cross-references: UNIPROT:P40859; GB:D16521; NID:G984787; PIDN:BAA03967.1; PID:G984788

C;Genetics:

A;Gene: sox

C;Superfamily: sarcosine oxidase

C;Keywords: oxidoreductase

Query Match 7.6%; Score 192; DB 2; Length 390;

Best Local Similarity 23.2%; Pred. No. 2.7e-07;

Matches 97; Conservative 68; Mismatches 162; Indels 92; Gaps 22;

QY 20 LLIVGAGTGTSTALHILARGYTNVTLDPYVPSPSAISAGNDVNKVISSGOYNNKDEIE 79

DB 7 VIVWGAGSMGMAGYQLAKQGVKTLV-DAPDPHTNGSHHGDTIRIHA-YGEGREYVP 64

QY 80 VNEILAEAFNGWKNDFPKPYVH-----DTGLMSACSQEGLDRLGVRVRPGEEDPNLVEL 135

DB 65 -----LALRQELWYE--LEKETHHKIFTKTGLV-----FGPK---GESAFVAET 105

QY 136 TRPEQFRKLAPGVLQGD-----FPGW---KGYYA--RSGAGWAHARNALVAARQAORM 185

DB 106 MEAAKEHSLTVD-LLEGDEINKRWPGITVPENYNAIFEPNSGVLFSENCIRAYRELAEAR 164

QY 186 GVKFTVGTPOGRVVTLLIPENNVDKGAVTGDKIWRARTFLCAGASAGQFLDFKNQLRPT 245

DB 165 GAKVLTHT---RVEDFDISPDVSK-IETANGS-YTADKLIVSNGAWNSKLLSKLN----- 214

QY 246 AWTLVHIALKPEERALKNPVFNIERGFFFPDEBERGEIKICDEHGPYTNMVQSADGT 305

DB 215 -----LDLPLQ-----YRQV-----GFFSEDSKYSNDI--DFPGF--MVEYPNGI 253

QY 306 MMSIPF-----EKTQIPKAEATRVALLKETMPQLADRPFSF 342

DB 254 YYGFPFSGGGLKGVHTFGQKIDPTINKRFVGYPDESNLRAFLLEYMPG-ANGELKR 312

QY 343 ARICWCADTANREFLIDHPQYHSLVLCGASGRGFKYLPISGNLIVD-AMEGKVPQKI 400

DB 313 GAVCMYTKLDEHFIILHPEHSNVVIAAGFSGHGFKFSSGVGEVLSQALTGKTBDHI 371

RESULT 4

JS0671

sarcosine oxidase (EC 1.5.3.1) precursor - Streptomyces sp.

C;Species: Streptomyces sp.

C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Jun-2000

C;Accession: JS0671; PS0345

R;Suzuki, K.; Ogishima, M.; Sugiyama, M.; Inouye, Y.; Nakamura, S.; Imamura, S. Bioeci. Biotechnol. Biochem. 56, 432-436, 1992

A;Title: Molecular cloning and expression of a Streptomyces sarcosine oxidase gene in St

A;Reference number: JS0671; MUID:92330009; PMID:1368326

A;Accession: JS0671

A;Molecule type: DNA

A;Residues: 1-389 <SUZ>

A;Cross-references: GB:D10623; NID:G217023; PIDN:BAA01473.1; PID:G217024

A;Experimental source: strain KB210-8S

A;Accession: PS0345

A;Molecule type: protein

A;Residues: 2-31;122-143;230-259;268-283 <SUZ1>

A;Note: the authors translated the initiation codon GTG for residue 1 as Val

C;Comment: This enzyme catalyzes the oxidation of sarcosine to form formaldehyde, glycine

C;Superfamily: sarcosine oxidase

C;Keywords: oxidoreductase

F;2-389/Product: sarcosine oxidase #status experimental <SAR>

F;11-16/Region: nucleotide binding #status predicted

Query Match 7.2%; Score 181; DB 2; Length 389;

Best Local Similarity 22.9%; Pred. No. 2e-06;

Matches 100; Conservative 59; Mismatches 150; Indels 128; Gaps 24;

QY 20 LLIVGAGTGTSTALHILARGYTNVTLDPYVPSPSAISAGNDVNKVISSGOYNNKDEIE 79

DB 7 VIVIGGMSGAAAHLSARG-ARVLGLEKF-----GPFVNRGSSHG 47

Qy 80 VNEILAEAFNGKNDPLKPYVHTDGLTMSACS-QSGLDRL-----GVRVPG 128
Db 48 GSRITRQSYF-----EDFAYVP-----LLRAYLYELEERATGRNVTATCGGVNAGPPD 97
Qy 129 DPNLVELTRP-----EQFRKLAPEGVLQDFFPGWKGYFARSAGWAHA 171
Db 98 SRTVSGSLRSATWDLAHEMLDAKEIRRRPPTLAPDDDEVALF-----EAKGLLRP 149
Qy 172 RNALVAAREAQRMGVKFTGTPTGGRVVTLLFENNVDKAVTGDG-KIWAERTFLCAGA 230
Db 150 ENMVAALHLQATROGAE-----LRFEEPLVRWBPYRDRGVHTGENTY----- 192
Qy 231 SAGQFLDKQLRPTANTLVHIALKPEERALKYKNIPIVFNIERG--FFEPDEBERGIKI 288
Db 193 TAGQLV-----ICPGAWA-----PO---LLADIGVPIITVERQIMWFQPKGTGPF-V 236
Qy 289 CDEHPGTNNVQSADGTMW-SIP-----FEKTO--IPKEAET-----RVRAL 327
Db 237 PERHPVY--IWEDADGVVYGFPAIDGPEKGAQVAFPRKQHTTPTETIDRTVHAHEVRAM 294
Qy 328 LKETMPQADRPFSF--ARICWCADTANREFLIDRHPQY-HSLVGLCGAGRGPKYLP 384
Db 295 ADHMSALIPDLGFTFLKAATCMYNTPDHEFVIARHPAHPESVTVACGFGSHGKFPV 354
Qy 385 GNLIWD-AMEGKVPQKI 400
Db 355 GEILADLALTGATAHPI 371

RESULT 5
JC7256

L-pipecolate oxidase (EC 1.-.-.-) - human
C;Species: Homo sapiens (man)
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C;Accession: JC7256
R;Jlalt, L.; de Kromme, I.; Oostheim, W.; Wanders, R.J.A.
Biochem. Biophys. Res. Commun. 270, 1101-1105, 2000
A;Title: Molecular cloning and expression of human L-pipecolate oxidase.
A;Reference number: JC7256
A;Accession: JC7256
A;Molecule type: mRNA
A;Residues: 1-390 <JL>
A;Cross-references: UNIPROT:Q9P029
C;Comment: This enzyme, a peroxisomal enzyme and a flavoprotein, which converts L-pipecolate to pipecolate.
C;Superspecies: UNIPROT:Q9P029
C;Keywords: PAD; flavoprotein; oxidoreductase

Query Match 7.0%; Score 177.5; DB 2; Length 390;
Best Local Similarity 21.7%; Pred. No. 3.7e-06;
Matches 86; Conservative 66; Mismatches 174; Indels 71; Gaps 16;
Qy 21 LIVGAGTGTSTALHLARRGTYNTVLDPPVPSAISAGNDVNVKVISSGOYNNKDEIEV 80
Db 11 IVIGAGLQGFETAYHLAKH-RKRILLLEQFLPHSRGSSHCQSRRIIKAYLE-----DF 63
Qy 81 NEILAEAFNGKNDPLKPYVHTDGLTMSACSQEGDLRLGVRVRPGEDPNLVELTRPEQ 140
Db 64 YTRMHCEYQIWAQ-----LEHEAGTQLH--RQTGLLLGKMKENQELKTIQANLSRQV 115
Qy 141 FRKLAPEGVLQDPP-----GWKGYFARSAGWAHAHARNALVA---AARE---AQRMGVK 188
Db 116 EHQLSSEELKQRPFNIRLPRGEVGLLDNSG-GVIYAYKALRALQDAIRQLGGVDRDEK 174
Qy 189 FVTGTPGQVRVTLFENNVDKAVTGDGKIWAERTFLCAGASQGLDFKNQLRPTAWT 248
Db 175 VVEINP-GLLV-----VKTTSRSTQAKSLVITAGWTNQL-----LRPLGIE 216
Qy 249 LVHIALKPEERALKYKNIPIVFNIERG-----FFEPDEBERGIKICDEHPG 294
Db 217 MPLQTLRINCYWREMPGSGVGSQAPPCFLWGLCPHHIYGLTGETVPGMLKVSYYHGN 276
Qy 295 YTNMVQSADGTMMSIPPEKTIQPKAEATRV-RALLKETMPQADRPFSFARICWCADTAN 353

Db 277 H-----ADPEERDCPTARTDI---GDVQLSSFVRDHLDPKPEP-AVIESCMYNTWTPD 326
Qy 354 REFLLDRHPQYHSLVLCGASGRGFKYLPISIGNLIWD 390
Db 327 EQFLDRHPKYDNIVIGAGFGSHGFKLAPVVGKILYE 363
RESULT 6
T44248
sarcosine oxidase (EC 1.5.3.1) [imported] - Arthrobacter sp. (strain TE1826)
C;Species: Arthrobacter sp.
A;Variety: strain TE1826
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 12-Jul-2004
C;Accession: T44248
R;Nishiya, Y.; Toda, A.; Imanaka, T.
Mol. Gen. Genet. 257, 581-586, 1998
A;Title: Gene cluster for creatinine degradation in Arthrobacter sp. TE1826.
A;Reference number: 222735; MUID:98223334; PMID:9563845
A;Accession: T44248
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-389 <NIS>
A;Cross-references: EMBL:AB007122; NID:g3116219; PIDN:BAA25926.1; PID:g3116221
A;Experimental source: strain TE1826
C;Keywords: oxidoreductase

Query Match 6.7%; Score 170; DB 2; Length 389;
Best Local Similarity 21.1%; Pred. No. 1.4e-05;
Matches 91; Conservative 69; Mismatches 171; Indels 100; Gaps 20;
Qy 12 LAVTKSSLLIVGAGTGTSTALHLARRGTYNTVLDPPVPSAISAGNDVNVKVISSQY 71
Db 1 MSIKQYDVIIVGCGSMGMAAGYLLSKQGVKTLIV-DSFDPPTNGSHHGDTIRIHA-Y 58
Qy 72 SNKQDEIEVNEILAEAFNGKNDPLKPYVHTDGLL-----MSACSOBGL 117
Db 59 GEGREYVFP-ALRAQELWYLEKETHHK-IFTKTGLVLFPGKGEAPFVAETMEAKHSL 116
Qy 118 DRLGVRVRPGEDPNLVELTRPEQFRKLAPEGVLQDFFGKMGYFARSAGWAHAHARNALVA 177
Db 117 D-----VDLSESEINKRWP-GVTYPE--NYNAIFEKN-SGVLFSENCIRA 158
Qy 178 AAREAQRMGVKFTGTPTGGRVVTLLFENNVDK-----GAVTGDGKIWAERTFLCAGASA 232
Db 159 YRELAEEANGAKVLTYTP---VEDFEIAEDFKIQTAYGSET-----ASKLIIVSMGAWN 208
Qy 233 GQFLDFKNQLRPTAWTLVHIALKPEERALKYKNIPIVFNIERGFFEPDEBERGE----- 285
Db 209 SKLSKLN-----IEIPLQP-----YRQVV-----GFFECDEKKYSNTHGYPA 246
Qy 286 --IKI-----CDEHPGTNNVQSADGTMMSIPPEKTIQPKAEATRVRLALKE 330
Db 247 FMVEVPTGIYVGPFGCGGLKIGYHTYGGKIDPDTINREF---GIYPEDSEGNIRKFL 303
Qy 331 TMPQADRPFSFARICWCADTANREFLIDRHPQYHSLVLCGASGRGFKYLPISIGNLIWD 390
Db 304 YMPG-ATGELKSGAVCMYTKTTPDEHFVIDLHPQSNVAIAAGFGSHGFKFSVVGETLSQ 362
Qy 391 -AMEGKVPQKI 400
Db 363 LAVTGKTEHDI 373

RESULT 7
B65058

fixC protein homolog b2766 - Escherichia coli (strain K-12)
N;Contains: probable guanine reductase (EC 1.6.5.-)
C;Species: Escherichia coli
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002
C;Accession: B65058
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-427 <STO>
A:CROSS-references: UNIPROT:Q91067; GB:AE004705; GB:AE004851; PIDN:AAG0616
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2776
C:Superfamily: hypothetical protein HI0499

Query Match 5.8%; Score 146.5; DB 2; Length 427;
Best Local Similarity 23.7%; Pred. No. 0.0012;
Matches 98; Conservative 59; Mismatches 164; Indels 93; Gaps 23;

Qy 18 SLLIVGAGTGTSTALHLARRGYNTVTLDPVPVPSAISAGNDVNKVKISSGGYSNNKDE 77
: : : ||| ||| ||| :
Db 30 TDVCVICAGVTGLSTALLFLENGF-KVTLEAAKVGFAS-GRNGGQIVNS--YSRDIDV 85
: : : ||| ||| ||| :
Qy 78 IEVN-----EILAEAFNGKNKDPLFKPYHDTGL-----LMSACSOEGLDRLGV 123
: : : ||| ||| ||| :
Db 86 IERTVGKREQLLGEMAFEGR---IIRVARVGIQCDLKOGGVFAAFTEQMDDLRAQ 142
: : : ||| ||| ||| :
Qy 124 VRPGE--DPNLVETRPQFRKLPAEVLQDFPGWKGYFARGSGAGWAHARNALVAARE 181
: : : | :
Db 143 KQLWERYGHNOLEIMDKARIR----EVATDNVIG--GMLDMSG-GHIHPLNALGEAAA 195
: : : | :
Qy 182 AQRMGVKFVTGTPQGRVTVLIFENNVDYKGAVTGDGKIWRARTFLCAGASAGOLFDFKNQ 241
: : : ||| ||| ||| :
Db 196 VESLGGRIYEQSPATRI-----ERGASPVVHTPOGV---KAKFIVVAGNA--YLN--- 241
: : : ||| ||| ||| :
Qy 242 LRPTAWTLVHIALKPERALYKNIP-----VI-----FNIERGFFPEDEERGETIKIDEHP 293
: : : ||| ||| ||| :
Db 242 -----GLVP ELAA--KSMPCGTQVITTEPLENEELAHSLLPQD-----YCVEDC 282
: : : ||| ||| ||| :
Qy 294 GYTNNVQSADCTWMSI-----PFKEQTQPKAEATRVALLKETMPOLADRPFSEARICWCA 349
: : : ||| ||| ||| :
Db 283 NYLDDYRLSDKKRLIYGGVVIYGARDPADIEAIRPMUKTFPQLKDVKIDFA---W-- 337
: : : ||| ||| ||| :
Qy 350 DTANREFLDRHPQ-----YHSLVLGCAGRGFKYLFPSIGNLIIVTDAMEGK 395
: : : ||| ||| ||| :
Db 338 -TGNFLLTLSELFPQVRIGDNIYYSQ---GCSGHGVYTHVAGKVLAEALRGQ 386
: : : ||| ||| ||| :

RESULT 13
S16133
dimethylglycine dehydrogenase (EC 1.5.99.2) - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S16133
R;Lang, H.; Polster, M.; Brandsch, R.
Eur. J. Biochem. 198, 793-799, 1991
A>Title: Rat liver dimethylglycine dehydrogenase. Flavinylolation of the enzyme in hepatocytes
A:Reference number: S16133; MUID:91266966; PMID:1710985
A:Accession: S16133
A:Molecule type: mRNA
A:Residues: 1-857 <LAN>
A:CROSS-references: UNIPROT:Q63342; EMBL:X55995; NID:X55995; PIDN:CAA39468.1; PID:g56689
C:Keywords: flavoprotein; oxidoreductase

Query Match 5.6%; Score 142.5; DB 2; Length 857;
Best Local Similarity 20.3%; Pred. No. 0.0065;
Matches 98; Conservative 74; Mismatches 160; Indels 151; Gaps 23;

Qy 21 LIVGAGTGTSTALHLARRGYNTVTLDPVPVPSAISAGN-----DYNKV 65
: : : ||| ||| ||| :
Db 46 VIIIGGCCVGSVLAYHLKAGMRDVLLE---KSELTAGSTWHAAGLTUFTPHGINLUKKI 101
: : : ||| ||| ||| :
Qy 66 ISSGOYSNNKDEIEVNEILAEEAFN--GWKNDLFPKPYHDTGLLSMASCOSQEGLDRLGV 123
: : : ||| ||| ||| :
Db 102 -----HYDSIKLYELEETQQVVG-----FHOPGSIRLATTPVERVEFYQ 143
: : : ||| ||| ||| :
Qy 124 V-RPGEDPNLVETRPQFRKLAP-----EGVLQDFFGKWKGYFARGSGAGWAHARNALVAA 178
: : : ||| ||| ||| :
Db 144 MTRTNWTHATEQYIIIEKHIELFPLNNMDKITLAGY-----NPQGDHIDPYSLWTAL 195
: : : ||| ||| ||| :

[illegible]

RESULT 15

Search completed: May 27, 2005, 13:48:47
Job time : 37.2303 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2005, 13:38:36 ; Search time 55.3518 Seconds
(without alignments)
975.171 Million cell updates/sec

Title: US-10-622-893A-5
Perfect score: 2530
Sequence: 1 MGSSGDDDLALAVTKSSSL.....LEGLPINPLRTGHHHHH 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 538906 seqs, 114359116 residues

Total number of hits satisfying chosen parameters: 538906

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New.*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US11_NEW_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	226.5	9.0	459	8	US-60-643-717-13558
2	181.5	7.2	440	7	US-11-097-143-4818
3	177.5	7.0	390	6	US-10-990-328A-13084
4	177.5	7.0	445	6	US-10-990-328A-13083
5	169	6.7	387	6	US-10-990-477-6
6	169	6.7	389	6	US-10-534-583-1
7	165	6.5	387	6	US-10-990-477-2
8	150	5.9	774	6	US-10-990-328A-13788
9	150	5.9	866	6	US-10-990-328A-13787
10	145.5	5.8	388	6	US-10-526-324-372
11	130.5	5.2	386	6	US-10-526-324-29
12	126.5	5.0	862	1	PCT-US05-10680-20
13	126.5	5.0	862	7	US-11-096-051-20
14	125.5	5.0	570	6	US-10-800-077-386
15	124.5	4.9	530	8	US-60-669-241-28797
16	119	4.7	437	7	US-11-040-686-33
17	119	4.7	437	7	US-11-040-686-31
18	119	4.7	437	7	US-11-040-686-37
19	117.5	4.6	374	6	US-10-924-025A-118
20	117.5	4.6	459	7	US-11-040-686-24
21	117.5	4.6	459	7	US-11-040-686-26
22	117.5	4.6	459	7	US-11-040-686-31
23	115.5	4.6	389	1	PCT-IB03-06509-1503
24	115	4.5	460	6	US-10-996-104A-22
25	115	4.5	505	7	US-11-031-175-9986

26	112.5	4.4	407	6	US-10-996-104A-20	Sequence 20, Appl
27	111.5	4.4	879	6	US-10-204-639-58	Sequence 58, Appl
28	109	4.3	42	6	US-10-532-053-15	Sequence 15, Appl
29	108.5	4.3	320	1	PCT-US05-04621-10	Sequence 10, Appl
30	108.5	4.3	320	7	US-11-057-027-10	Sequence 10, Appl
31	108	4.3	796	6	US-10-450-763-48529	Sequence 48529, A
32	107.5	4.2	270	8	US-60-645-354-36	Sequence 36, Appl
33	107	4.2	421	7	US-11-040-686-42	Sequence 42, Appl
34	106.5	4.2	822	8	US-60-643-717-17883	Sequence 17883, A
35	105.5	4.2	296	1	PCT-US05-10680-18	Sequence 18, Appl
36	105.5	4.2	296	7	US-11-096-051-18	Sequence 18, Appl
37	105.5	4.2	822	8	US-60-643-717-7212	Sequence 7212, Ap
38	105	4.2	1664	6	US-10-450-763-44587	Sequence 44587, A
39	104.5	4.1	211	8	US-60-645-354-44	Sequence 44, Appl
40	104.5	4.1	339	7	US-11-122-943-10	Sequence 10, Appl
41	104.5	4.1	339	7	US-11-122-943-12	Sequence 12, Appl
42	104	4.1	273	8	US-60-645-354-40	Sequence 40, Appl
43	104	4.1	308	1	PCT-US05-02996-6	Sequence 6, Appl
44	104	4.1	308	7	US-11-046-644-6	Sequence 6, Appl
45	104	4.1	515	7	US-11-097-143-42765	Sequence 42765, A

ALIGNMENTS

RESULT 1
US-60-643-717-13558
; Sequence 13558, Application US/60643717
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)A
; CURRENT APPLICATION NUMBER: US/60/643,717
; CURRENT FILING DATE: 2005-01-12
; NUMBER OF SEQ ID NOS: 19247
; SEQ ID NO 13558
; LENGTH: 459
; TYPE: PRT
; ORGANISM: ASPERGILLUS NIDULANS FGSC A4
; US-60-643-717-13558

Query Match 9.0%; Score 226.5; DB 8; Length 459;
Best Local Similarity 19.8%; Pred. No. 1.2e-10;
Matches 90; Conservative 38; Mismatches 101; Indels 225; Gaps 11;

Qy	15	TKSSSLIVGAGTGTSTALHARRGYTNTVLD-PYPVPSAISAGDNVKNVSSGQYSN	73
Db	5	SKSIPAIIGDAGFLSTALHLVQNGYDTITVLEQDEKIPPPYSAANYLNKIV-----	57
Qy	74	NKDEIEVNEILABEAFNGWKNNDPLFKPYHDTGLLMSACSQEGLDRLGVRVRPGEDPNLV	133
Db	58	-REBY-----EDP---	64
Qy	134	ELTRPEQPRKLAPEGVLQDPPGKGYFARSGAGWAHARNALVAAAREAQRMGVKFTGT	193
Db	65	-----	64
Qy	194	PQGRVVTLI FENNNDKGA VTDGDKI WRAERTFLCAGASAGQFLDFKNQLRPTAWLVHIA	253
Db	65	-----	64
Qy	254	LKPEERALYKNI PVIFNI ERGFFPEPDERGEIKICDEHPGYTNMVQSADGTMMSIPEK	313
Db	65	-----LCPMGSGYIN-TDKATGVSHS-PFPV	95
Qy	314	TQ-----IPKEATRVRLIKETMPOLADRPSPFAICWCADTANFEFLDRHP-QVHS	366
Db	96	NOEASSGFLPAEDETRIRKLLQOTLPALARNPLVLKSLCWPADTKDSDFIIDFVPGSKGS	155
Qy	367	LVLGCGASGRGFKYLPISIGNLIVDAMEGKVPQKTHELIKWNPDI AANNRWRDITLGRFGP	426
Db	156	WVIESADSGHGFQWFFIVGWSVTSLLSKRHTPNSIH---SGKPNLRANMGQSTYKNTGLP	212

```
QY 427 NRVMDPHDVKEWTVN-----QYRDISKLKGELEG 455
Db 213 WVLPSVQARRGFENKGLVHEYLPIRLKGLREG 246

RESULT 2
US-11-097-143-4818
; Sequence 4818, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4818
; LENGTH: 440
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-4818

Query Match 7.2%; Score 181.5; DB 7; Length 440;
Best Local Similarity 24.1%; Pred. No. 6.7e-07;
Matches 108; Conservative 57; Mismatches 149; Indels 135; Gaps 24;

QY 17 SSSILVAGTGTSTALHARRGY-----TNTVLDPPVPVPSAISAG-NDVNKVISSG- 69
Db 31 SCGLVIIGGGMGASSAFWLKSRALQLGRKLNVLVVER-----DAGYTSASTVLSVGG 83

QY 70 --QVSNKDEIEVNEILAEAFNGWKNDPLFKPYHYDHTGLLMSACSOEGL----- 117
Db 84 VRQOFSLAENIEMSLFGYNFVNGREHLGDVLDLCYQPNGLILA-SEKGAHILAKNSKLQ 142

QY 118 DRIGVRVRPGEDPNLVLTREPRK-----LAPEGVLOG-----DFPGWKGYFARSGAGWA 169
Db 143 NELGAR-----NELLGPEALRQRFPWLSTEGVELGCHGIDKEGWF-----PWA 186

QY 170 HARNALVAAREAQRMGVKFTGTPQGRVVTILFENND-----VKGAVTGDGKIWAERTF- 225
Db 187 -----LLMGYKKARALGANFANGSVWG-----PEWNDGSLGAVVDAGDVLQRTVKFD 236

QY 226 ---LCASAGQ---FLDDNQLRPTATWLTVHIALKPEERALLY-----KNIPVI---FN 270
Db 237 TCVLGAYSGQVARLAGIDGKEAKESLSVALPVERKRYVVVSTQGNCRGLATPLT 296

QY 271 IE-RGFFPEPD-----EERGEIKICDEHPGTYTNMVQSADGTMSIPPEKTQ 315
Db 297 VDPDGTFRDRGLCGNFCRSPNEDEPECETILDVDHG-----FE--- 338

QY 316 IPKEAETRVALLKETWPLQADRPFSFARI-----CWCA-----DTANREFLIDRHPQVHSL 367
Db 339 -----TDVWPTLANRVAFESVKIOSSWAGFYDHTNTFDANGVIGRHPHYSNL 385
```

```
QY 368 VLGCASGRGFKYLPISIGNLIVD-AMEGK 395
Db 386 FIAAGFSGHGIQOTPAVGRAISLILDGK 414

RESULT 3
US-10-990-328A-13084
; Sequence 13084, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13084
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-13084

Query Match 7.0%; Score 177.5; DB 6; Length 390;
Best Local Similarity 21.7%; Pred. No. 1.2e-06;
Matches 86; Conservative 66; Mismatches 174; Indels 71; Gaps 16;

QY 21 LIVGAGTGTSTALHARRGYTNVTVDPPVPVPSAISAGNDVNKVISSGQYSNNKKEIEV 80
Db 11 IVIGAGIQGCTAYHLAKH-RKRILLLEQFFLPHSRGSSHGQSRIIRKAYLE-----DF 63

QY 81 NEILAEAFNGWKNDPLFKPYHYDHTGLLMSACSOEGLDRLGLVRVRPGEDPNLVLTREPEQ 140
Db 64 YTRMMECYQIWAQ-----LEHAECTQLH--RQTGLLLGMKENQELTKIQANLSRQRV 115

QY 141 FRKLAPEGVLOGDFF-----GMKGYFARSGAGWAHARNALVA---AARE---AORMGVK 188
Db 116 EHQLSSEELKQRFNIRLPRGEVGLDMSG-GVIYAKALRALQDAIRQLGGIVRDEK 174

QY 189 FVTGTPQGRVVTILFENNDVKGAVTGDGKIWAERTFLCAGASAGQFLDKNQLRPTAWT 248
Db 175 VVEINP-GLLVT-----VKTTSRYQAKSLVITAGPWTNQL-----LRPLGIE 216

QY 249 LVHIALKPEERALLYKNIPVIFNIERGEF-----FFEPDEERGEIKICDEHPG 294
Db 217 MPLQTLRINVCYWRENVPGSYGVSAFPCEFMLGLCFPHIYGLTGEYPLGMKVSYHHGN 276

QY 295 YTNMVQSADGTMSIPPEKTQIPKEAETRV-RALLKETWPLQADRPFSFARI-CWCA DTAN 353
Db 277 H-----ADPEERDCPTARDI---GDVQILSSFVRDHLPLDKPEP-AVIESCMYTNTPD 326

QY 354 REFLIDRHPQVHSLVLCGASGRGFKYLPISIGNLIVD 390
Db 327 EQFLIDRHPKYDNIVTIGAGFSGHGFKLAPVVGKILYE 363

RESULT 4
US-10-990-328A-13083
; Sequence 13083, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13083
; LENGTH: 445
; TYPE: PRT
```

```

; ORGANISM: Homo sapiens
US-10-990-328A-13083

Query Match          7.0%; Score 177.5; DB 6; Length 445;
Best Local Similarity 21.7%; Pred. No. 1.5e-06;
Matches 86; Conservative 66; Mismatches 174; Indels 71; Gaps 16;

Qy 21 LIVAGTGTGTATLHLARRGYNTVTLDPYPVPSAISAGNDVNKVISSSQYNNKDEIEV 80
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 IVIGAGTGCCTAYHLAKH-RKRILLLEQPFPHSRGSGHQSRIRKAYLE-----DF 118
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 81 NEILAEAFNGKNDPLFKPYTHDTGLLMSACQEGDLRLGVRVPCGEDPNLVLTPEQ 140
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 YTRMHCEYQIWAQ-----LEHAGQLH--ROTGLLLGKMNENBELKTIQANLSKQRV 170
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 141 FRKLAPEGVLQGDPP-----GWKGYPARSGAGWAHARNALVA---AARE---AQRMGVK 188
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 EHQCLSSEELKQFPNNIRLPRGEVGLDNGS-GVIYAKALRALQDAIRQLGGVDRGEK 229
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 189 FVTGTPGGRVVUTLIFENNVDKAGVTGDKTWRARTPLCAGASAGQPLDFPNQLRPTAWT 248
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 230 VWEINP-GLLVT-----VKTTSRYQAQSLVITAGPWTNQL-----LRPLGIE 271
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 249 LVHIALKPEERALYKNIPVFNTERGP-----FPEPDEERGEIKICDEHPG 294
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 272 MPLQTLRINVCYWRMVPGSYGVSAQPCFLWLGLCPHHIYGLPTGYPLGMKVSYHHGN 331
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 295 YTNMVSQADGTMMSIPPEKTOIPKEAETRV-RALLKETMPQLADRPSPFARICWCA DTAN 353
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 332 H-----ADPEERDCPTARDI---GDVOILSSFVRDHLDPDLKPEP-AVIESCMYTNTPD 381
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 354 REFLDRHPQYHSLVLGCGASGRGFKYLPISGNLIVD 390
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 382 EQFLDRHPKYDNIVIGAGFSGHGFKLAPVVGKILYE 418
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
US-10-990-477-6
; Sequence 6, Application US/10990477
; GENERAL INFORMATION:
; APPLICANT: KIKKOMAN CORPORATION
; APPLICANT: FURUKAWA, Keisuke
; APPLICANT: KAJIYAMA, Naoki
; TITLE OF INVENTION: MODIFIED SARCOSINE OXIDASES, MODIFIED SARCOSINE OXIDASE GENES,
; TITLE OF INVENTION: AND METHODS FOR PREPARING THE MODIFIED SARCOSINE OXIDASES
; FILE REFERENCE: 261714USO
; CURRENT APPLICATION NUMBER: US/10/990,477
; CURRENT FILING DATE: 2004-11-18
; PRIOR APPLICATION NUMBER: JP 2003-387975
; PRIOR FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: JP 2004-184960
; PRIOR FILING DATE: 2004-06-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Bacillus sp. strain KS-11A
US-10-990-477-6

Query Match          6.7%; Score 169; DB 6; Length 387;
Best Local Similarity 21.4%; Pred. No. 6.4e-06;
Matches 89; Conservative 69; Mismatches 172; Indels 86; Gaps 20;

Qy 20 LLIVAGTGTGTATLHLARRGYNTVTLDPYPVPSAISAGNDVNKVISSSQYNNKDEIE 79
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7  VIVVGAGSMGMAAGYTLAKQGVKTLV-DAFDPHTEGSHHGDTRIIRHA-YGEGREYVP 64
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 80 VNEILAEAFNGKNDPLFKPYTHDTGLL-----MSACSQEGL-----D 118
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 F-ALRAQELMYELENETHNK-IFTKTVGVFGPKGESDFVAETMEAAAEHSLTVDLLEGD 122
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 119 RLGVRRVPRGDDPNLVLTPEQFRKLAPEGVLQGDFFPGWKGYFARSAGAGWAHARNALVAA 178

```

Db	123	Eintr-wpg-----itvpenynai-----fepnsgvlfsencirsy 157
Qy	179	Areaqrmgvkfvtgtpqgrvvltli fenndvkgavtdgdkimwraertflcagasaqofldf 238
Db	158	relavakagakillyt---rvedfvsodqv-k-1qtangs-ytadkllvsmgawnskllsk 212
Qy	239	knqlrptawtlvhiaikp-----beralykn---ipv-ifniergffpepderge 285
Db	213	ln-----ldlplqyrqvqvwffpdsneakysndvypafmvevpkgyiycgfpfsgg- 262
Qy	286	ikicdehpcytnmvqsadgtmmsipektoi pkeatrvrallkettmqoladrpfspari 345
Db	263	----cglkigyhygqoidpdtinrfegayq---edesnlrdelekympe-angelkrgav 315
Qy	346	cwcaadtaneflidrhpoyslvlgcgsgrgfkylpsicnlivd-amegkvpqi 400
Db	316	cmvtyktpdehfidthpeshnvpvaagfsgghgfkfssvvgvlsqatlattgktehdi 371
RESULT 6		
US-10-534-583-1		
: Sequence 1. Application US/10534583		
: GENERAL INFORMATION:		
: APPLICANT: KISHIMOTO, Takahide		
: APPLICANT: SOGABE, Ateushi		
: APPLICANT: OKA, Masanori		
: TITLE OF INVENTION: MODIFIED SARCOSINE OXIDASE, PROCESS FOR PRODUCING THE SAME AND		
: TITLE OF INVENTION: REAGENT COMPOSITION USING THE SAME		
: FILE REFERENCE: 235670		
: CURRENT APPLICATION NUMBER: US/10/534,583		
: CURRENT FILING DATE: 2005-05-11		
: PRIOR APPLICATION NUMBER: PCT/JP2003/014423		
: PRIOR FILING DATE: 2003-11-13		
: PRIOR APPLICATION NUMBER: JP2002-329427		
: PRIOR FILING DATE: 2002-11-13		
: PRIOR APPLICATION NUMBER: JP2002-329428		
: PRIOR FILING DATE: 2002-11-13		
: PRIOR APPLICATION NUMBER: JP2003-33641		
: PRIOR FILING DATE: 2003-02-12		
: NUMBER OF SEQ ID NOS: 14		
: SOFTWARE: PatentIn version 3.1		
: SEQ ID NO 1		
: LENGTH: 389		
: TYPE: PRT		
: ORGANISM: Arthrobacter SP. TE1826		
US-10-534-583-1		
Query Match 6.7%; Score 169; DB 6; Length 389;		
Best Local Similarity 21.1%; Pred.No. 6.4e-06;		
Matches 91; Conservative 69; Mismatches 171; Indels 100; Gaps 20;		
Qy	12	LatvkssslilvgactwctstalhlarqytnvtvldpvpvsaisagndvknvissgoY 71
Db	1	MSIKKDYDIVVVGASGMGAAGYLLSKQGVKTLV--DSPHPHTNGSHHGDTRIIRHA-Y 58
Qy	72	SNNKDEIYNEIILAEAFNGKNDPLFKPYHYDHTGLL-----MSACSQEGL 117
Db	59	GEGREYVFP-ALRAQELWYLEKEETHK-IFTKYGLVFGPKGEAPFVAETMEAAKHSU 116
Qy	118	DRLGVVRPGEDPNLVLPTRPQFRKLAPGVLGQDFGKMGYFARSAGAHARNALVA 177
Db	117	D-----VDLLEGSEINKRP-GVTVPE--NNAIFEKN-SGVLFSENCIRA 158
Qy	178	AREAQRMGVKFTGTGQGRVVLTli fenndvkgavtdgdkimwraertflcagasa 232
Db	159	YRELAEPANGAKVLTYP---VEDFEIADFVKIQTAYGSFT-----ASKLIVSMGAWN 208
Qy	233	GQFLDFPKNLQRPATWTLVHIAIKPBERALYKNIPVIFNIERGFFPEPDERGE----- 285
Db	209	SKLLSKLN-----IEIPLQ-----YRQV-----GFFECDEKKYNTHGYP 246
Qy	286	--IKI-----CDEHPGYTNMVQSADGTMMISPFPEKTOIPKEATRVRLAKE 330

```
Db 247 FMVEVPTGIYGPSPGGCLKIGYHTYQKIDPDTINREP---GIYPEDEGNIRKFLT 303
Qy 331 TWPOLADRPFPSPARICWCADTANREFLIDRHPQYHSLVLCGASGRGFKYLPISIGNLIYD 390
Db 304 YMG-ATGELKSGAVCMYTKTPDEHFVIDLHPQFSNVAAGSFSGHGFSSVVGTELQ 362
Qy 391 -AMEGKVPQKI 400
Db 363 LAVTGKTEHDI 373

RESULT 7
US-10-990-477-2
; Sequence 2, Application US/10990477
; GENERAL INFORMATION:
; APPLICANT: KIKKOWAN CORPORATION
; APPLICANT: FURUKAWA, Keisuke
; APPLICANT: KAJIYAMA, Naoki
; TITLE OF INVENTION: MODIFIED SARCOSINE OXIDASES, MODIFIED SARCOSINE OXIDASE GENES,
; TITLE OF INVENTION: AND METHODS FOR PREPARING THE MODIFIED SARCOSINE OXIDASES
; FILE REFERENCE: 261714US0
; CURRENT APPLICATION NUMBER: US/10/990,477
; CURRENT FILING DATE: 2004-11-18
; PRIOR APPLICATION NUMBER: JP 2003-387975
; PRIOR FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: JP 2004-184960
; PRIOR FILING DATE: 2004-06-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-990-477-2

Query Match 6.5%; Score 165; DB 6; Length 387;
Best Local Similarity 21.4%; Pred. No. 1.4e-05;
Matches 89; Conservative 68; Mismatches 173; Indels 86; Gaps 20;

Qy 20 LLIVGAGTGTSTALHARRGYNVTVDLPYPVPSAISAGNDVKNVSSQYQNNKDEIE 79
Db 7 VIVVGAGSMGWAAGYLLAKGVKTLV-DADPPHTGSHGDTRIIRHA-YGEGREYVP 64
Qy 80 VNEILAEAFNGKNDPLFPYVHDTGLL-----MSACSQBGL-----D 118
Db 65 F-ALRAQELWYLENETHNK-IFTKTGLVFGPKGESDFVAETMEAAAEHSLIVDLLEG 122
Qy 119 RLGVVRPQGEDPNLVELTRPEQFKLAPEGLQDPFGWKGYFARSGAGWAHARNALVAA 178
Db 123 EINTR-WPG-----ITVPENYNAI-----PEPNSGVLFSENCIRSY 157
Qy 179 AREAQRMGKVFVTGTPQGRVVTLIFENNVDKGAVTGDKIWRARTPLCAGASAGQFLDF 238
Db 158 RELAVAKGAKLITY---RVDEPVSQDVK-IQTANGS-YTADKLIVSGAWNSKLSK 212
Qy 239 KNQLRPTAWTLVHIALKP-----EERALYKN-----IPV-IFNIERGFFFPDEERGE 285
Db 213 LN-----LDIPLQYQVRVGVGFFDSNEAKYSNDVDYPAFMVVEVPKGIYYGFPSPFG- 262
Qy 286 IKTCDEHPGYTNVQSGADGTMMSIPFEKTOIPKEAETRVALLKETMPQADRPFPSPARI 345
Db 263 ---CGLKIGYHTYQKIDPDTINREFGAYQ---EDESNLRDFLBKYMPE-ANGELKKGAA 315
Qy 346 CWCADTANREFLIDRHPQYHSLVLCGASGRGFKYLPISIGNLIYD-AMEGKVPQKI 400
Db 316 CMYTKTPDEHFVIDTHEPSNVFVAAGSFSGHGFSSVVGVELSQLATTGKTEHDI 371

RESULT 8
US-10-990-328A-13788
```

```
; Sequence 13788, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13788
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-13788

Query Match 5.9%; Score 150; DB 6; Length 774;
Best Local Similarity 20.8%; Pred. No. 0.00066;
Matches 102; Conservative 77; Mismatches 187; Indels 124; Gaps 23;

Qy 2 GSGGDDDDLALAVTK---SSSLIVGAGTGTSTALHARRGYNVTVDLPYPVPSAISA 58
Db 31 GREGEKPPLSAETQWKDRAETVILGGCGVSLAYHLAKAGKDVLLVLE---KSELTA 86
Qy 59 GN-----DYNKVISSQYQNNKDEIEVNEILAEAFN--GWKNDPLFKPY 101
Db 87 GSTWHAAGLTTFPHPGINLKKI-----HYDSIKLYEKLBEETQGVVG-- 128
Qy 102 YHDTGLLMSACSOEGLDRLGVRV-RPEDPNLVELTRPEQFKLAP-----EGVLCQDFPG 156
Db 129 FHQPGSIRLATTVPVRVDEFKYQWTRTGMHATEQYLLIEPEKIOEMFFLLNNKVLGLY-- 186
Qy 157 WKGYPARSAGWAHARNALVAAAREAQRMGVKFTVGTPOGRVVTLIFENNVDKGAVTG 216
Db 187 -----NPGDGHIDPYSLTWALAAGARKCGALLKYAP---VTSLKASDGTWVETPOG 237
Qy 217 KIWAERTFLCAG---ASAGQFLDFKNQLRPTAWTLVHIALKPEERALYKNIPVFNIE- 272
Db 238 SM-RANRIVNAAGFWAREVGKMGLEHPLIPVQHQQVVTSTIPEVKALKRELPLVLRDLEG 296
Qy 273 -----RGFPFPEPDEERGEIKIDE-----HPCGYTNV--QSADGTMMSIPFEKTOI 316
Db 297 SYYLQERDGLLFGPYESQEKMKVQDSWVTNGVPPGFKELFESDLDRIMEHI----- 349
Qy 317 PKEAETRVALLKETMPQADRPFPSPARICWCADTANREFLIDRHPQYHSLVLCGASGR 376
Db 350 -KAAMENVPVKKADIINVNGPITYS-----PDILP---WVCPHOGVRNYWAIG-FGY 399
Qy 377 GFKYLPSIGNLIYD-AMEGKVPQKIHILIKWNPDIANRNWRTLGRFGGPNRMDFH 435
Db 400 GIIHAGGVGKGLSDWILLHGEPP---FDLIEDPN-----RYG----- 433
Qy 436 KEWNTVOYRD 445
Db 434 -KWTITQYTE 442

RESULT 9
US-10-990-328A-13787
; Sequence 13787, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13787
; LENGTH: 866
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-13787

Query Match          5.9%; Score 150; DB 6; Length 866;
Best Local Similarity 20.0%; Pred. No. 0.00077;
Matches 102; Conservative 77; Mismatches 187; Indels 124; Gaps 23;

Qy 2 GGGGDDDLALAVTK-----SSLLIVGAGTGTSTALHLARRGYNTVLDPPYVPVSAISA 58
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 31 GREGEKPLPSAETQWKDRAETVIGGCGVSLAYHLAKAGMDVVLL-----KSELTA 86

Qy 59 GN-----DVKNVSSGOYNNKDEIEVNEILAEAFN--GKNDKPLPKY 101
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 87 GSTWHAAGLTYFHPGINKLKI-----HYDSIKLYEKELEBETGOVVG----- 128

Qy 102 YHDTGLMSACSGELDRGVRV-RPGEDPNLVELTRPEQFKLAP-----EGVLQGFPG 156
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 FHQGSIRLATTVPVRVDEPKYQMTRTGHWATEQYLIIEPKIQEMFPLNNKVLGLY-- 186

Qy 157 WKGYFARSGAGWAHARNALVAAREAQRMGVKFTGTGPGRVTTLIFENNDRKGAVTGDG 216
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 187 -----NPGDGHDPYSLTWALAAGARKCGALLKYPAP---VTSLKARSDGTWDVETPQG 237

Qy 217 KIWAERTFLCAG---ASAGQFLDFKNQLRPTATWLVIHIALKPERALYKNIPVFNIE- 272
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 SM-RANRIVNAAGFWAREVGMKMGLEHPLIPVQHQYVVTSTIPEVKALKRELPLVRLDEG 296

Qy 273 -----RGFFPDEERGEIKICDE-----HPGYTNMV--QSADGTWMSIPFEKTOI 316
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 SYVLQERDGLLPGPYEQSKMKVQDSWVTNGVPPGKGELPESDLDRIMEHI----- 349

Qy 317 PKEATRVRALLKETPQLADRPFSFARICWCADTANREFLIDRHPQYHSLVLGCGASGR 376
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 350 -KAAMEMVPLKXADIINVNGPITYS-----PDILP---MVGPHQGVRYNVAIG-FGY 399

Qy 377 GFKYLPSIGNLVD-AMEGKVPQKIHELIIKNPDIANRWRDTLGRFGGPNRVMDFHDV 435
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 400 GIHAGGVGKYLSDWILHGPP---FDLIEDPN-----RYG----- 433

Qy 436 KEWTVQYRD 445
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 434 -KWTITQYTE 442

RESULT 10
US-10-526-324-372
; Sequence 372, Application US/10526324
; GENERAL INFORMATION:
; APPLICANT: Imanaka, Takayuki
; APPLICANT: Atomi, Haruyuki
; TITLE OF INVENTION: METHOD OF TARGETED GENE DISRUPTION, GENOME OF
; TITLE OF INVENTION: HYPERTHERMOSTABLE BACTERIUM AND GENOME CHIP USING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 490051.401USPC
; CURRENT APPLICATION NUMBER: US/10/526,324
; CURRENT FILING DATE: 2005-02-28
; PRIOR FILING DATE: PCT/IB2003/003597
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: JP 2002-319011
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 2167
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 372
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Thermococcus kodakaraensis KOD1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (414542)..(414542)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (786890)..(786890)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (786907)..(786907)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (786944)..(786944)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (839139)..(839139)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1128488)..(1128488)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1128499)..(1128499)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1128505)..(1128505)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1128517)..(1128518)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1128539)..(1128540)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1561400)..(1561400)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1561477)..(1561477)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1561545)..(1561545)
; OTHER INFORMATION: n is a or c or g or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1767941)..(1767941)
; OTHER INFORMATION: n is a or c or g or t.
US-10-526-324-372

Query Match          5.8%; Score 145.5; DB 6; Length 388;
Best Local Similarity 20.0%; Pred. No. 0.00061;
Matches 99; Conservative 64; Mismatches 167; Indels 165; Gaps 20;

Qy 12 LAVTKSSSLIVGAGTGTSTALHLARRGYNTVLDPP-YVPVSAISAGNDVNVKVISGQ 70
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 MTWKSEAKTVIIGGGIIGLSIAYNLAKLGESDIVLEKGY-----LNGSTFRCTGTI 54

Qy 71 YSNKDEIEVNEI-LAEAEFNGWKNDDPLFKPYHDTGLMSACSGELD-----R 119
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 55 RQOFGEANIRMMKRSVELMKLKEELGYDVEFTQSGYLEFLIYSEEBLEAFNNVRLQNR 114

Qy 120 LGYVRPGEDPNLVELTRPEQFKLAP-----EGVLQGFPGWKGYFARSGAGWAHAR--- 172
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 115 FGVPFR-----IITPEEAKIIVPLPLNTDGI-----AAMNHTDGA 151

Qy 173 ---NALVAAAREAQRMGVKFTGTGPGRVTTLIFENNDRKGAVTGDGKIWAERTFLCAG 229
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 152 NPFAKVPAYANAARKLGVIEIYET---EAKDIKVEDGKIKAVVTNRGEI-RTGRVINAAN 207

Qy 230 ASAGQFLDFKNQLRPTATWLVIHIALKPERALYKNIPVFNIERGFPEDEERGEIK-- 287
```



```
; PRIOR APPLICATION NUMBER: US To be assigned
; PRIOR FILING DATE: 2005-03-30
; PRIOR APPLICATION NUMBER: 60/557,978
; PRIOR FILING DATE: 2004-03-30
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 20
; LENGTH: 862
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-10680-20

Query Match      5.0%; Score 126.5; DB 1; Length 862;
Best Local Similarity 21.3%; Pred. No. 0.073;
Matches 71; Conservative 34; Mismatches 103; Indels 125; Gaps 15;

Qy 196 GRVVTLIFENNVDKGVGTGDKIWRARTFLCAGASAGQFLDFKNQLRPTAWTLVHIALK 255
Db 599 GASLTILFERSPP---LTQYHTVWIPWNVF-----YVMDTLVMK 634
Qy 256 PEER-----ALYKNIPVIFNIERGEFFPEDEERGEIKICDEHPGYTNMVQSGDTMM 307
Db 635 KEENDIPSCDLGPFVRNPPIVSSPLSTFFRSPSPEDSPI----- 673
Qy 308 SIP-----FEKTOIP-----KEAETRVALLKETMPQLADRPSPFARI----- 345
Db 674 -IPETQVLHEETTIPGTDLKLSYLSSRAAGYSVLKITMTQ-SIIPFNLMKVLHMLVAVVG 731
Qy 346 ----CWCADTAN--REFLIDRHPQYHSLVLCGAS--GRGPKYLPISIGNLIVDAMEGKVP 397
Db 732 RLFOKWPFPASNLAYTFIWDKTDAYNKQYGLSEAVSVGYEY----- 774
Qy 398 QKIHELKWNPDIAANRNWRDTLGRFGGPNRVMDFHDVKWTVNQYRDISKLKGE----- 452
Db 775 ESCLDLTWEKRTAILQGYELDASNMGG--WTLDKHHV---LDVQNGILYKNGENQFIS 829
Qy 453 -----LEGLPINPLL-----RTGHHHHH 472
Db 830 QPPVWSLEGKPIPNPLGLDSTRTGHHHHH 862

RESULT 14
US-10-800-077-386
; Sequence 386, Application US/108000077
; GENERAL INFORMATION:
; APPLICANT: Reddy, Ramachandra
; APPLICANT: Gill, Parkash
; TITLE OF INVENTION: NUCLEIC ACID COMPOUNDS FOR INHIBITING
; FILE REFERENCE: VASG-P01-001
; CURRENT APPLICATION NUMBER: US/10/800,077
; PRIOR FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: US 60/454,432
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/454,300
; NUMBER OF SEQ ID NOS: 396
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 386
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Recombinant B4ECv3 protein
US-10-800-077-386

Query Match      5.0%; Score 125.5; DB 6; Length 570;
Best Local Similarity 23.0%; Pred. No. 0.05;
Matches 49; Conservative 22; Mismatches 63; Indels 79; Gaps 11;

Qy 295 YTNMVQSGADGTM-----SIPFEKTOIPKEAETRVALLKETMPQLAD-----RPFSFAR 344
Db 402 YTFEVTALNGVSSLATGVPFPEPVNVTTD-----REVPPAVSDIRVTRSPSSUS- 451
Qy 345 ICWCADTANREFLIDRHPQYHSLVLCGASG-----RGFKYLPIS 384
Db 452 LANAVPRAPSGAWLDYEVKYEK---GARGPSSVRFLKTSNPAELRGLKRGASLYLVQV 507
Qy 385 GNLIVDAMEGKVPQKITHELIKWNPDIAANRNWRDTLGRFGGPNRVMDFHDVKWTVNQYR 444
Db 508 -RARSEAGYGPFQGEHHS---QTQLDESEGWRREQ-----GSKRAIL----- 544
Qy 445 DISKLKGELEGLPIPNPLL-----RTGHHHHH 472
Db 545 -----QIEGKPIPNPLGLDSTRTGHHHHH 570

RESULT 15
US-60-669-241-28797
; Sequence 28797, Application US/60669241

Qy 196 GRVVTLIFENNVDKGVGTGDKIWRARTFLCAGASAGQFLDFKNQLRPTAWTLVHIALK 255
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2005, 21:41:14 ; Search time 921 Seconds
(without alignments)
9459.321 Million cell updates/sec

Title: US-10-622-893A-6

Perfect score: 1419

Sequence: 1 atgggaggttcgggtgacga.....atccaccatcatcatcattaa 1419

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5700845 seqs, 3069779757 residues

Total number of hits satisfying chosen parameters: 11401690

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1419	100.0	1419	19	US-10-622-893A-6
2	335.2	23.6	1320	19	US-10-470-678-1
3	333.6	23.5	1320	19	US-10-470-678-3
4	332	23.4	1320	19	US-10-470-678-2
5	332	23.4	1320	19	US-10-470-678-4
6	62.8	4.4	1314	16	US-10-232-655-4
7	57	4.0	422	18	US-10-767-701-22052
8	50.4	3.6	1314	16	US-10-232-655-2
9	38.8	2.7	3731	17	US-10-600-230-78
10	37.2	2.6	1302	9	US-09-815-242-9647
11	37	2.6	951	17	US-10-282-122A-25526
					Sequence 25526, A
					Sequence 6, Appli
					Sequence 1, Appli
					Sequence 3, Appli
					Sequence 2, Appli
					Sequence 4, Appli
					Sequence 4, Appli
					Sequence 22052, A
					Sequence 2, Appli
					Sequence 78, Appli
					Sequence 9647, Ap

12	36.6	2.6	660	14	US-10-123-155-480	Sequence 480, App
13	36.6	2.6	660	15	US-10-146-731-480	Sequence 480, App
14	36.6	2.6	660	15	US-10-140-472-480	Sequence 480, App
15	36.6	2.6	660	15	US-10-141-761-480	Sequence 480, App
16	36.6	2.6	660	16	US-10-142-885-480	Sequence 480, App
17	36.6	2.6	660	16	US-10-158-790-480	Sequence 480, App
18	36.6	2.6	660	17	US-10-137-871-480	Sequence 480, App
19	36.6	2.6	660	17	US-10-140-923-480	Sequence 480, App
20	36.6	2.6	660	17	US-10-141-756-480	Sequence 480, App
21	36.6	2.6	660	17	US-10-141-759-480	Sequence 480, App
22	36.6	2.6	660	17	US-10-140-805-480	Sequence 480, App
23	36.6	2.6	660	17	US-10-140-864-480	Sequence 480, App
24	36.6	2.6	660	17	US-10-142-426-480	Sequence 480, App
25	36.6	2.6	671	14	US-10-184-644-346	Sequence 346, App
26	36.2	2.6	671	14	US-10-184-634-346	Sequence 346, App
27	36.2	2.6	2634	18	US-10-437-963-18493	Sequence 18493, A
28	36	2.5	2113	18	US-10-425-115-182534	Sequence 182534, A
29	35.8	2.5	739	14	US-10-123-155-478	Sequence 478, App
30	35.8	2.5	739	15	US-10-146-731-478	Sequence 478, App
31	35.8	2.5	739	15	US-10-140-472-478	Sequence 478, App
32	35.8	2.5	739	15	US-10-141-761-478	Sequence 478, App
33	35.8	2.5	739	16	US-10-142-885-478	Sequence 478, App
34	35.8	2.5	739	16	US-10-158-790-478	Sequence 478, App
35	35.8	2.5	739	17	US-10-137-871-478	Sequence 478, App
36	35.8	2.5	739	17	US-10-140-923-478	Sequence 478, App
37	35.8	2.5	739	17	US-10-141-756-478	Sequence 478, App
38	35.8	2.5	739	17	US-10-141-759-478	Sequence 478, App
39	35.8	2.5	739	17	US-10-140-805-478	Sequence 478, App
40	35.8	2.5	739	17	US-10-140-864-478	Sequence 478, App
41	35.8	2.5	739	17	US-10-142-426-478	Sequence 478, App
42	35.6	2.5	414	10	US-09-918-995-33483	Sequence 33483, A
43	35.4	2.5	3659	18	US-10-437-963-2456	Sequence 2456, Ap
44	35.2	2.5	882	14	US-10-184-644-574	Sequence 574, App
45	35.2	2.5	882	14	US-10-184-634-574	Sequence 574, App

ALIGNMENTS

RESULT 1

US-10-622-893A-6
; Sequence 6, Application US/10622893A
; Publication No. US20050014935A1
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong-Sheng
; APPLICANT: Datta, Abhijit
; APPLICANT: Wang, Yuping
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: DETERMINATION OF GLYCATED PROTEINS
; FILE REFERENCE: 46692001300
; CURRENT APPLICATION NUMBER: US/10/622,893A
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1419
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding a chimeric protein
US-10-622-893A-6

Query Match	100.0%;	Score 1419;	DB 19;	Length 1419;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1419;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATCGGAGGTTCCGGTGACGATGATGACCTCGCTCGCCCTCACTAGTCAATCTCTC	60	
Db	1	ATCGGAGGTTCCGGTGACGATGATGACCTCGCTCGCCCTCACTAGTCAATCTCTC	60	
Qy	61	CTGATCTTGGTCCGGGACTTGGGGACCTCAACGGCTCTGCACCTCGCGCGCGGGA	120	
Db	61	CTGATCTTGGTCCGGGACTTGGGGACCTCAACGGCTCTGCACCTCGCGCGCGGGA	120	

QY 121 TATACCAACGTTACCGTGTGGACCCCTATCTGTCTAGCCCATCTCCGCGGAAC 180
DB 121 TATACCAACGTTACCGTGTGGACCCCTATCTGTCTAGCCCATCTCCGCGGAAC 180
QY 181 GAGTGAACAAAGTCATTAGCATGGCCAAATATTCGAATAACAAAGACGAATCGAAGTG 240
DB 181 GAGTGAACAAAGTCATTAGCATGGCCAAATATTCGAATAACAAAGACGAATCGAAGTG 240
QY 241 AATGAGATCTTGGCGAAGAGCGTTTAAGGTTGGAAGAACCAACCGCTTTTCAACCG 300
DB 241 AATGAGATCTTGGCGAAGAGCGTTTAAGGTTGGAAGAACCAACCGCTTTTCAACCG 300
QY 301 TATTATCATGATACGGGCCCTGCTGATGTCGTCTGCTCGAGAGGCGCTGAGTCGCTG 360
DB 301 TATTATCATGATACGGGCCCTGCTGATGTCGTCTGCTCGAGAGGCGCTGAGTCGCTG 360
QY 361 GCGTCCGGGTACGTCGGCGGAGGATCTTAATCTGGTGAACTTACCCGCGCGAGCAA 420
DB 361 GCGTCCGGGTACGTCGGCGGAGGATCTTAATCTGGTGAACTTACCCGCGCGAGCAA 420
QY 421 TTTCTGTAACCTGCGCGGAGCGGTGTCAGAGGTGATTTTCCGGGTTGGAAGGGTAC 480
DB 421 TTTCTGTAACCTGCGCGGAGCGGTGTCAGAGGTGATTTTCCGGGTTGGAAGGGTAC 480
QY 481 TTTGCGGCTTCCGCGCTGGCTGGGCACATGCAAGGAATGCCCTTAGTGGCAGCAGCACG 540
DB 481 TTTGCGGCTTCCGCGCTGGCTGGGCACATGCAAGGAATGCCCTTAGTGGCAGCAGCACG 540
QY 541 GAAGCAGCGCATGGGTGTAATTTGTTGTCGCAACCCCGCAGGGTGGTGTAGTCACG 600
DB 541 GAAGCAGCGCATGGGTGTAATTTGTTGTCGCAACCCCGCAGGGTGGTGTAGTCACG 600
QY 601 TTAATCTTTGAAATACAGATGTAAGGTCGCTTACCGGCGATGCGAAATTTGCGA 660
DB 601 TTAATCTTTGAAATACAGATGTAAGGTCGCTTACCGGCGATGCGAAATTTGCGA 660
QY 661 GCGGAACGTACATCTCTGTGTCTGGGCTAGCGCGGTGAGTTCCTAGATTTCAAGAT 720
DB 661 GCGGAACGTACATCTCTGTGTCTGGGCTAGCGCGGTGAGTTCCTAGATTTCAAGAT 720
QY 721 CAATCTCGACCAACCGCTTGGACCCCTGGTACATTCGCTGTTAAACCGGAAGAACGTGCG 780
DB 721 CAATCTCGACCAACCGCTTGGACCCCTGGTACATTCGCTGTTAAACCGGAAGAACGTGCG 780
QY 781 TTGTCAGAAATATACCGGTTATCTTTAATCGAAGCGGGGTTTCTTGTGAACCGAT 840
DB 781 TTGTCAGAAATATACCGGTTATCTTTAATCGAAGCGGGGTTTCTTGTGAACCGAT 840
QY 841 GAGGAGCGCGGTGAGATTAAATATGCGATGAACACCCCGGCTACACAAATATGTTCCAG 900
DB 841 GAGGAGCGCGGTGAGATTAAATATGCGATGAACACCCCGGCTACACAAATATGTTCCAG 900
QY 901 AGTGCAGACGCAACGATGAGATTCCTGTTGAAAAAAACCCAGATTCGAAAGAACGC 960
DB 901 AGTGCAGACGCAACGATGAGATTCCTGTTGAAAAAAACCCAGATTCGAAAGAACGC 960
QY 961 GAAACGCGGCTTCCGGCCCTGCTGAAGAGACAAATGCCCCAGCTGGCAGACCGTCCATTC 1020
DB 961 GAAACGCGGCTTCCGGCCCTGCTGAAGAGACAAATGCCCCAGCTGGCAGACCGTCCATTC 1020
QY 1021 AGCTTGCAGCAATTTGCTGGTGGCGATACCGCAATCGCGAATTCCTGATAGATCGA 1080
DB 1021 AGCTTGCAGCAATTTGCTGGTGGCGATACCGCAATCGCGAATTCCTGATAGATCGA 1080
QY 1081 CATCCGAGTACCAAGTCTTGTGTTGGGCTGTTGGTGGCGAGGAGAGGGTTTAAATAT 1140
DB 1081 CATCCGAGTACCAAGTCTTGTGTTGGGCTGTTGGTGGCGAGGAGAGGGTTTAAATAT 1140
QY 1141 CTGCTTCTATTGGGAATCTCATTTGACCGCATGGAAGTAAAGTCCGCGCAAAATTT 1200
DB 1141 CTGCTTCTATTGGGAATCTCATTTGACCGCATGGAAGTAAAGTCCGCGCAAAATTT 1200

QY 1201 CACGAATTAATCAAGTGGAAACCCGGACATTGGCGGAAACCGTAATGGCGTGATCTCTG 1260
DB 1201 CACGAATTAATCAAGTGGAAACCCGGACATTGGCGGAAACCGTAATGGCGTGATCTCTG 1260
QY 1261 GGGCGTTTGGCGGTCCAAATCGTGTGATGGATTTTCATGATGTGAAGTAAGCAAT 1320
DB 1261 GGGCGTTTGGCGGTCCAAATCGTGTGATGGATTTTCATGATGTGAAGTAAGCAAT 1320
QY 1321 GTTCAGTATCGTGATATTTTCCAAAGCTGAAAGGAGAGTTCGAAAGTAAAGCAATCCCTAAC 1380
DB 1321 GTTCAGTATCGTGATATTTTCCAAAGCTGAAAGGAGAGTTCGAAAGTAAAGCAATCCCTAAC 1380
QY 1381 CCGTTACTCGGCACAGGCCATCACCATCATCATTA 1419
DB 1381 CCGTTACTCGGCACAGGCCATCACCATCATCATTA 1419
RESULT 2
US-10-470-678-1
; Sequence 1, Application US/10470678
; Publication No. US20050101771A1
; GENERAL INFORMATION:
; APPLICANT: KOUZUMA, Takuji et al.
; TITLE OF INVENTION: COMPOSITION FOR ASSAYING GLYCATED PROTEINS
; FILE REFERENCE: 1516-0121P
; CURRENT APPLICATION NUMBER: US/10/470,678
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: JP 2001/ 22953
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: JP 2001/ 39796
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: JP 2001/ 24002
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1320
; TYPE: DNA
; ORGANISM: Fusarium oxysporum IFO-9722
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1320)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: DDBJ E16562
; DATABASE ENTRY DATE: 1999-07-28
; PATENT DOCUMENT NUMBER: JP 1998201473-A/1
; PATENT FILING DATE: 1997-01-20
; PUBLICATION DATE: 1998-08-04
US-10-470-678-1
Query Match 23.6%; Score 335.2; DB 19; Length 1320;
Best Local Similarity 55.3%; Pred. No. 3.4e-99;
Matches 721; Conservative 0; Mismatches 568; Indels 15; Gaps 3;
QY 41 TCATTAAGTCATCATCTCTCTGATCGTTGGTGGCGGACCTTGGGGACCTTCAACGGCTC 100
DB 11 TCACCAACACAGTCCCAAAATTCATCGTTGGTGGCGGAACTTGGGGATGCTCAATGCC 70
QY 101 TGCACCTCGCGCGCCGGGATATACCAAGTTACCGTGTGAGACCCCTATCTGTCCCTA 160
DB 71 TCCATCTCGCCGCTCGGGGTACACCAACGTCATCTGTTCTCGATGTCAATTCGATCCCGT 130
QY 161 GCGCATCTCCGCGGAAACGAGCTGAACAAAGTCATTAGCAGTGGCAATATTCGAATA 220
DB 131 CACCGATATCAGCGGGGCTGATGTAAACAACTTCTGCTGGCCGACTGTGCACTGCCGATA 190
QY 221 ACAAA-----GACGAAATCGAAGTGAATGAGATCTTGGCGGAAGAGCGTTTAAACGGTT 274
DB 191 GCAAAAGTGTATGATGAAGACTCAATCTGGAAAGCACTTAGCTACGCCGAGCTCAAGGAT 250
QY 275 GGAAGAACACCGCGCTTTTCAAAACCGTATATCATGATACGGGCCCTGCTGATGCTGCTT 334
DB 251 GGCTCCACGACCCCTGTCTTCCAAACCATCTTGCCCAACATACAGGCTCTGTGCTGGTGGCT 310

Qy	635	TTACGGCGGATGGCAAAATTTGGAGAGCGGNA	CGTCAATTCCTCTGTGTGCTGGGGCTACGC	694
Db	611	AGACGGCAGATGGTAAGGACACAGACGGATCGAACTA	TTCTTTCGCTGGTGTCTTCAG	670
Qy	695	CGGTCAGTTCCTAGATTTCAAGAATCAACTTTCGACCAACCGCTTGGACCCCTGGTACACA	754	
Db	671	CAGAGTTCTTCCTCGATTTTCAGAAACAGATCCAGCCCTACGGCGTGGACCTTGGGCCATA	730	
Qy	755	TTGCGTTAAAAACCGGAAGACGTGCGTTGTACAAAAATATACCGGTTATCTTTAAACATCG	814	
Db	731	TCCAGATGACACACGAAGAAACCAAGCTGTACAGAAACCTGCCACCTCTTTTCAACATCA	790	
Qy	815	AACGGGGTTTTTCTTTTGAACCCGATCAGAGCGCGTGAAGATTAAATAATGCGATCAAC	874	
Db	791	ACCAAGTTTCTTCATGGAACCTGATGAGATCTTTCATCACTCAAGATGTGCGATGAAC	850	
Qy	875	ACCGGGCTACACAATATATGGTCCAGAGTGCAGACGGCACG---	931	
Db	851	ATCCGGGCTACTGCAACTGGGTTGAAAAGCTGGTTTCTAAGTAGTACCCCGAGTCCATCCCT	910	
Qy	932	TCGAAAAAACCCAGATTCCAAAGAAAGCCGAACCGCGGTTTCGGGCCCTGCTGAAAGAGA	991	
Db	911	TCGCAAAAGCATCAAGTGCCAACCGAGGCTGAACGACGCATGAAGCAGTCTCTGAAAAGATA	970	
Qy	992	CAATGCCCCAGCTGGCAGACCGTCCATTTCAGCTTTCGCACGCATTTCGTGTGTGTCGCCGATA	1051	
Db	971	TCATGCTCAGCTTGACATCGGCCGCTTGTTCATGCTCGAACTCTGCTGGTGGCTGATA	1030	
Qy	1052	CCGGGAATCGGGAATTCCTGATAGATCGACATCCGGCAGTAGCA	CAGTCTGTGTGTGTGGGCT	1111
Db	1031	CACAGSATGAATGTTCTCTGATCACTATCATCTTCGACATCCCTCACTTGTTCATTCCTT	1090	
Qy	1112	GTGTGCGAGCGGAAGAGGGTTTAAATATCTGCCCTTCTATTGGGAATCTCAATTGTTCAGC	1171	
Db	1091	CAGTGATTTCGGGCAACGGGTTACATGATATCAATCAATTCGAAAGTTCATCTCTGACT	1150	
Qy	1172	CGATGGAAGTAAAGTCCGCAAAAAATTCAGAAATTAATCAAGTGGAAACCCGGACATTC	1231	
Db	1151	GTATGAGGGGATACGCTTTGAGGAAAGGTTTGCCAAAGTTCTCGAGATGCGCACAGAGAAGT	1210	
Qy	1232	CGGGCAACCGTAACTGCGGTGATACTCTGGGCGCTTTTGGCGGTCCAAATCTGTGTATGG	1291	
Db	1211	TTACCGAGTTCGGGGTAAGATCTCTGGATCGGTTTGGAGCTGACGATAAGTCAATGG	1270	
Qy	1292	ATTTTC-----ATGATGTGAAGAAATGGACCAATGTTTCAGTAT	1329	
Db	1271	ATTTTCCCAAGAGTGTAGTAGGGATGGCAAAATATCAAGAAT	1314	

RESULT 4

```

US-10-470-678-2
; Sequence 2, Application US/10470678
; Publication No. US2005010771A1
; GENERAL INFORMATION:
; APPLICANT: KOUZUNA, Takuji et al.
; TITLE OF INVENTION: COMPOSITION FOR ASSAYING GLYCATED PROTEINS
; FILE REFERENCE: 1516-0121P
; CURRENT APPLICATION NUMBER: US/10/470,678
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: JP 2001/ 22953
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: JP 2001/ 39796
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: JP 2001/ 24002
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence derived from Fusarium oxysporumIFO-9722

```

```
Db      911  TCGAAGAGCATCAAGTGCACCGAGGCTGAAACGCGATGAAGCAGTTTCTGAAAGATA 970
Qy      992  CAATGCCAGCTGGCAGACCGTCCATTACAGTTTCGACGCAATTTGCTGGTGTCGCCGATA 1051
Db      971  TCATGCCCTCAGCTTGACAGTCGGCGCTTGTTCTATGCTCGAATCTGCTGGTGCGCTGATA 1030
Qy     1052  CCGCGAATCCGAATTCCTGTATAGATCGACATCCGAGTACCACAGTCTTGTGTGGCT 1111
Db     1031  CACAGGATAGAATGTTCTGTATACCTATCATCTCGACATCCCTCACTGTCTATGCTT 1090
Qy     1112  GTGTGCGAGCGGAAGAGGGTTTAAATATATCTGCTTCTATTGGGAATCTCATTTGTGACG 1171
Db     1091  CAGGTGATTGGGCACCGGTTACTGGCATATCATCAATTTGGAAGTTTATCTCTGACT 1150
Qy     1172  CGATGGAAGTAAAGTCCCGCAAAAATTCACGAATTAATCAAGTGGAAACCCGACATTTG 1231
Db     1151  GTATGGAGGTTACGCTTGAGGAAGGTTTGCAAGTTCTGGAGATGGCGACGAGAAAT 1210
Qy     1232  CGCGAACCCTTAAGTGGGTTGATCTCTGGGCGTTTGGCGGTCGAAATCGTGTGATGG 1291
Db     1211  TTACCGAGTTCTGGGTTAAAGATCCTCTGGATCGGTTTGGAGCTGACGATAAGATCATGG 1270
Qy     1292  ATTTC-----ATGATGTGAAGAAATGGACCAATGTTTCAGTAT 1329
Db     1271  ATTTGCCAAGATGATGTAGAGGATGGACAAATATCAAGAAT 1314
```

RESULT 5

US-10-470-678-4

; Sequence 4, Application US/10470678

; Publication No. US20050101771A1

; GENERAL INFORMATION:

; APPLICANT: KOZUUMA, Takuji et al.

; TITLE OF INVENTION: COMPOSITION FOR ASSAYING GLYCATED PROTEINS

; FILE REFERENCE: 1516-0121P

; CURRENT APPLICATION NUMBER: US/10/470,678

; CURRENT FILING DATE: 2003-07-30

; PRIOR APPLICATION NUMBER: JP 2001/ 22953

; PRIOR FILING DATE: 2001-01-13

; PRIOR APPLICATION NUMBER: JP 2001/ 39796

; PRIOR FILING DATE: 2001-02-16

; PRIOR APPLICATION NUMBER: JP 2001/ 24002

; PRIOR FILING DATE: 2001-08-08

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 4

; LENGTH: 1320

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: sequence derived from Fusarium oxysporum IFO-9722

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1320)

; FEATURE:

; NAME/KEY: mutation

; LOCATION: (1114)..(1115)

US-10-470-678-4

```
Query Match      23.4%; Score 332; DB 19; Length 1320;
Best Local Similarity 55.1%; Pred. No. 3,9e-98;
Matches 719; Conservative 0; Mismatches 570; Indels 15; Gaps 3;
```

```
Qy      41  TCATAGTCATCATCTCTCTGTATCGTTGGTCCGGGACTTGGGGCACTTCAACGGCTC 100
Db      11  TCACCAACAGTCCCAAAATTCATCGTTGGTGGCGAACTTGGGGATGCTCAACTGCC 70
Qy     101  TGCACCTCGCGCGCGCGGATATACCAACGTTACCGTGGGACCCCTATCTGTCCTTA 160
Db      71  TCCATCTCGCGCGTGGGGTTTACCAACGTCACCTGTTCTCGATGTCAATCGCATCCG 130
Qy     161  GCGCCATCTCGCGCGGAAACGACGTGAACAAAGTCATTAGACGTGGCCAAATATTCGA 220
```

```
Db     131  CACGATATCAGCGCGGCATGATGTAAACAACAACTTGTCTGCCGCACTGTGCACTGCCGATA 190
Qy     221  AAAAA-----GACGAAATCGAAGTGAATGAGATCTTTGGCGGAGAGCGGTTTAAACGGTT 274
Db     191  GCAAGGTGATGATGAAGACTCAATCTGGAAGACACTTAGCTACGCCGAGCTCAAGGAT 250
Qy     275  GGAAGAACGACCCGCTTTTCAAACCGTATTATCATGATCGGGCTCTGCTGATGTCTGCTT 334
Db     251  GGCTCCACGACCCCTGTCTTCCAAACCAATTCTGCCCAATACAGGCTCTGTCTGGCTGGCT 310
Qy     335  GCTCGCAGAGGCGCTCGATCGCTGGCGCTCGGGTACGTCGCGGGCAGGATCCTTAATC 394
Db     311  CAACCAAAAGTCTATCAACGAGCTGGTAGAAGATGAGATCGGTGACGACATCGACCAGT 370
Qy     395  TGGTGGAACTTACCCGCCCGAGCAATTTTCGTAAACTTGGCCCCCGGAAGGGGTTGTCAAG 454
Db     371  ATACACCTCTCAACACACGACGAGAAGATTTTCAGAAAGACCATGCTCGAGGGTATCCTGACAG 430
Qy     455  GTGATTTTCCGGGTTGGAAGGGTACTTTTCGGCGGTTCCGGCGCTGGCTGGGCACATGCAA 514
Db     431  GTAACTTTCCAGGCTGGAGGGCTTTTACAAGCCACGCGGTTCTGGTTGGGTTTCATGCTC 490
Qy     515  GGAATGCTTAGTGGCAGCAGCAGCGAAGCACAGCGCATGGGTGTAATAATTTCTTACTG 574
Db     491  GAAAGCTATGAAAGCTGCTTTTCGAGAGAGCGAGAGGCTTGGTGTCAAATTTCACTACTG 550
Qy     575  GCACCCCGCAGGGTCTGTAGTCACTTAATCTTTGAAAAATAACGATGTAAAAAGTGGCG 634
Db     551  GCTCTCCGGAAGGAAGGTGGAGAGTCTGATCTTTTGAAGACGCGCATGTTTCGAGGTGCCA 610
Qy     635  TTAGCGCGCATGGCAAAATTTGGAGAGCGGAAAGTACATTTCTCTGTGTGCTGGGCTAGCG 694
Db     611  AGACGCGCATGGTGAAGGAGCAGACAGCGGATCGAACTATTTTTCGCGTGGTCTTCAG 670
Qy     695  CGGTCACTTCTAGATTTTCAAGAAATCAACTTTCGACCAACCGTGTGACCTGTGTACACA 754
Db     671  CAGAGTTCTTCTCGATTTTGAGAACCGATTCAGCTTGGCTGGACCTGGGCGCATATA 730
Qy     755  TTGCGTTAAACCGGAAGAACGTCGTTGTGTAACAAAATAATACCGGTTTATCTTTAACTCG 814
Db     731  TCCAGATGACACCAAGAAACCAAGCTGTACAAGAACCTGCCACCTCTTTTCAACATCA 790
Qy     815  AACGGGGGTTTTCCTTTGAACCCGATGAGAGAGCGGCTGAGATTAATAATATGCGATGAAC 874
Db     791  ACCAAGGTTTCTTCATGGAAACCTGATGAGGATCTTCACTCAACTCAAGATGCGATGAAC 850
Qy     875  ACCGGGCTACACAAATATGCTCCAGATGTCAGACGCGCAGC---ATGATGAGCATTCGCT 931
Db     851  ATCCGGGCTACTGCAACTGGGTTGAAAAGCCTGGTTCTAAGTACCCCGCAGTCCCATCCCT 910
Qy     932  TCGAAAAAACCCAGATTTCCAAAAAGAACCGCAACCGCGCTTCGGGCGCTCTGTAAGAGA 991
Db     911  TCGCAAGCATCAAGTGCACCGGCTGNAACGACGATGAAGCAGTTTCTGAAAGATA 970
Qy     992  CAATGCCAGCTGGCAGACCGTCCATTACAGTTTCGACGCGATTTGTGTGGTGCCGATA 1051
Db     971  TCATGCTCAGCTTGCAGATCGGCGCTTGTTCATGCTCGAAATCTGCTGGTGCGCTGATA 1030
Qy     1052  CCGCGAATCGGAATTCCTGTATAGATCGATCCGAGTACCACAGTCTTGTGTGGCT 1111
Db     1031  CACAGGATAGAATGTTTCTCTGATCACCTTATCATCTCGACATCCCTCACTGTCTATGCTT 1090
Qy     1112  GTGTGCGAGCGGAAGAGGGTTTAAATATATCTGCTTCTATTGGGAATCTCATTTGTGACG 1171
Db     1091  CAGGTGATTGGGCACCGGTTTACATCATATCAATCAATTTGGAAGTTTCACTCTGACT 1150
Qy     1172  CGATGGAAGTAAAGTCCCGCAAAAATTCAGAAATTAATCAAGTGGAAACCCGACATTTG 1231
Db     1151  GTATGGAGGTTACGCTTGAGGAAGGTTTGCCTCAAGTTCTGGAGATGGCGACGAGAAAT 1210
Qy     1232  CGCGAACCCTTAAGTGGGCTGATCTCTGGGCGGTTTGGCGGTCGAAATCGTGTGATGG 1291
```

Query Match	4.48;	Score 62.8;	DB 16;	Length 1314;
Best Local Similarity	48.2;	Pred. No. 3.2e-09;		
Matches 298;	Conservative 0;	Mismatches 287;	Indels 33;	Gaps 3;
QY	674	TCCTGCTGCTGGGGCTAGCGGGTCAGTTCCTAGATTTCAAGAAATCAACTTCGACCAA	733	
DB	650	TCCTGGCTGCTGGTGGCTGGAGTCCCACTTGGTGGATCTAGAAGATCAGTGTGTTTCAA	709	
QY	734	CCGCTTGGACCCCTGGTACACATTCGCTTAAACCCGGAAGAACGTGCGTGTGTACAAAATA	793	
DB	710	AGSCCTGGGTTTTCGCTCATATTCAACTCACACCCAAAGAGCGGACGTACAAAGAATG	769	
QY	794	TACGGTTATCTTTAAACATCGAACCGGGGTTTCTTTTGAACCCGATGAGGAGCGCGGTG	853	
DB	770	TGCGTGTGGTCTATGATGGTGAATATGGGTTCTTTTGTGACCCCAACGAGTA---TG	826	
QY	854	AGATTAATAATATGCGATGAACACCCGGGCTACACAA-----ATATGTGCC	898	
DB	827	TGATCAAGTCTGTGACGAGTTCCTCGTTTCTCTCGCTTCAAACTGCATCAACCGTAGC	886	
QY	899	AGAGTCGACAGCGGCAGCATGATGAGCATTCGGTTTCGAAAAACCCAGATTCCTCAAAGAAG	958	
DB	887	GGGCTGCATCTCCCAAGATGATATCCGTACGGCGATCACGCCCAAGCATCCCAACAGATA	946	
QY	959	-----CCGAAACGGCGGTTTCGGGGCCCTGTGTGAAGAGACAAATGCCCCAGC	1003	
DB	947	CCTACCTGTATGCTCCGGAAGTCAACCATAGCAAAAGCGATCGCAAGGTTTCCTGCGAGAAT	1006	
QY	1004	TGGCAGACCGTCCATTACGCTTTCGACGCAATTCCTGGTGGCGGATACCGCGAATCGCG	1063	
DB	1007	TTAAAGACAAGGAGCTCTTCAACCGTACCATGTGCTGGTGTACAGATACCGGCCGATGCTA	1066	
QY	1064	AAATTCCTGTAGATCGACATCCGACGATACCAACAGTCTTGTGTTCGGGCTGTGGTCCGAGCG	1123	
DB	1067	ACTTATTGATTTGGCAACACCCGAAGTGGAGAATTTCAATTCTCGCCCACTGGAGATAGCG	1126	
QY	1124	GAGAGGGTTTAAATATCTGCGCTTCTATTTGGGAATCTATTTGTGACCGCGATGGAAGGTA	1183	
DB	1127	GCAATTCCTTCAAGCTGTGCAAAACATCGGGAACACGTTGTTGAGCTTTTACGAGGAT	1186	


```
; FILE REFERENCE: 227590US0
; CURRENT APPLICATION NUMBER: US/10/232,655
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: JP 2001-266665
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: JP 2001-378151
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: JP 2002-228727
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Coniochaeta sp.
US-10-232-655-2

Query Match      3.6%; Score 50.4; DB 16; Length 1314;
Best Local Similarity 52.3%; Pred. No. 4e-05;
Matches 136; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

Qy 628 GTGTCCTTACGGCGGATGCGAAATTTGGAGCGGAACGTACATTCCTGTGTGCTGGG 687
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
604 GCGCTTGAGACGGCAGATGGTACCAATATTACCTGACAAGGTGGTCTTAGCAGCTGGC 663
Qy 688 GCTAGCGGGGTGAGTTCCTAGATTTCAAGAAATCACTTCGACCAACGGCTTGACCCCTG 747
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
664 GCATGGAGCCCAACCCCTGGTGGACCTGGAAGATCAATGTTGCTCGAAGGGCTTGGGTGTAT 723
Qy 748 GTACACATTCGTTTAAACCGGAAGACGTGCGTGTGTACAAAATATACCGGTTATCTTT 807
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
724 GCTCATATTGATGAGCGCTGAAGAGCGCGCTGAGTATAGGTGTCCAGTTGTGTAT 783
Qy 808 AACATCGAAGCGGGTTTTCTTTTGAACCGCATGAGGAGCGCGGTGAGATTTAAATATGC 867
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
784 AATGGCAATTTGGCTTCTCTTTGAGCCTAAT--GAGTTTGTGTGTAATAAAGGTGTGC 840
Qy 868 GATGAACACCGGGCTACAC 887
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
841 GACGAATTCGAGGATTC 860

RESULT 9
US-10-600-230-78/c
; Sequence 78, Application US/10600230
; Publication No. US20040092020A1
; GENERAL INFORMATION:
; APPLICANT: Wilkinson, Jack
; APPLICANT: McBride, Kevin
; APPLICANT: Bertain, Sean
; TITLE OF INVENTION: GENETIC CONSTRUCTS HAVING HETEROLOGOUS
; TITLE OF INVENTION: 3' POLYADENYLATION SIGNAL SEQUENCE MOTIFS THAT FUNCTION IN
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 0325.210
; CURRENT APPLICATION NUMBER: US/10/600,230
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: 60/390,529
; PRIOR FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 3731
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
US-10-600-230-78

Query Match      2.7%; Score 38.8; DB 17; Length 3731;
Best Local Similarity 46.6%; Pred. No. 0.48;
Matches 124; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

Qy 1061 GCGAATTCCTGATAGATCGACATCCGACGATACACAGTCTTGTGTGGGCTGTGGTSCGA 1120
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3297 GCGACTTCATATACCAACCATCCATCGCACACCTCGTCTCTTGCAACTGGTGCA 3238
```

```
Qy 1121 GCGAAGAGGGTTTAAATATCTGCTTCTAATTGGGAATCTCATTTGTTGACCGGATGGAAG 1180
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3237 GCGGCCACGCATACAAATTTCTCCCACTTATCGGCGAAAAAGGTGTCGATGCGCTGGAGG 3178
Qy 1181 GTAAAGTGCCGCAAAAAAATTCAGGAATTAATCAAGTGGAAACCGGACATTTGCGGCGGAACC 1240
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3177 GTAAAGTCCGACCCGGAACCTGCGGCGCTCTGGGACTGCGCGGAGGCTGTGGACAGAGGCC 3118
Qy 1241 GTAACTGGCGTGATCTCTGCGGCGTTTTTGGCGGTCCAAATCGTGTGATGGAATTTTCATG 1300
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3117 ACTTTGACGGAACAGAGGATGGAAGTCGTTTCAGGGGATAAGCGGCTGCTGCTTTTGGATG 3058
Qy 1301 ATGTGAAGGAATGACCAATGTTTCAG 1326
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3057 AGTTGCGGAAGGGGCGAGAAAGCGAAG 3032

RESULT 10
US-09-815-242-9647
; Sequence 9647, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9647
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1302)
US-09-815-242-9647

Query Match      2.6%; Score 37.2; DB 9; Length 1302;
Best Local Similarity 52.6%; Pred. No. 0.92;
Matches 81; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
```

```
Qy 31 GCTCTCGCGTCACTAAGTCATCATCTCTCTGATCGTGTGCGCGGACTTGGGGCACC 90
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 GCTCTGCAAGACCATATTGATGCGGATGTGGTATCATTCGCGGTGCTTTTCAGGCATC 126
Qy 91 TCAACGCGCTCTGCACCTCGCGCGCCCGGATATACCAACGTTACCGTGTGCGGCCCTAT 150
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
127 AATACCGCGCTGGAATCTGGCGGAGGAGGATCACTAATGTTGTTGTTCTTGGAGCCGCG 186
```

QY 151 CCTGTCCTTAGCGCCATCTCCGCGGGAACGACG 184
Db 187 CATCTGGGTATGGCGCACCGCGGAACGGCG 220

RESULT 11

US-10-282-122A-25526

; Sequence 25526, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 25526

; LENGTH: 951

; TYPE: DNA

; ORGANISM: Mycobacterium avium

US-10-282-122A-25526

Query Match 2.6%; Score 37; DB 17; Length 951;
Best Local Similarity 64.7%; Pred. No. 0.9;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 59 TCTGTGTCGTGGTGGCGACCTTGGGCACTTGGCGACCTCAACGGCTCTGCACCTCGCGCGCGG 118
Db 5 TCGTGTGTCATCGTGGCGGGGTGTGCGCTCGCGCGCGCCACGAGCTCTCGCGCGCGG 64

QY 119 GATATACCAAGTTACCGTCTGGA 143

Db 65 GACACGACGACGTGGTGTGCTGGA 89

RESULT 12

US-10-123-155-480

; Sequence 480, Application US/10123155

; Publication No. US20030068794A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-480

Query Match 2.6%; Score 36.6; DB 14; Length 660;
Best Local Similarity 8.5%; Pred. No. 1;
Matches 53; Conservative 193; Mismatches 372; Indels 3; Gaps 1;

QY 173 CCGAAACACGTGAACAAGTCATTAGCAGTGGCCAATATTGATAAACAAGACGAAA 232
Db 39 CQTGAFLVLLVTVIIVNIKLIDTRRAISEANEDPEQDYDEALGLEPPRRRSGPVRV 98
QY 233 TCGAAGTGAATGAGATCTTGGCGGAAGAGCGGTTAAACGGTTGGAAGAAGACACCGCTTT 292
Db 99 LDVEVYSSRSKVYVAVDGTTVLEDEAREQGRGHVILVQATGHVMAKRVFDYISHEDE 158
QY 293 TCAACCGGTATTATCATGATAGCGGCTGCTGATGTCTGCTGCTCGCAGGAGGCGCTGG 352
Db 159 AMVLFLNMVAPGRVLCTVKGEGSFLKDTAKALLRSLGSGAQAGPALGWRDTPAFVGRKGG 218
QY 353 ATGCGCTGGCGTCCGGGTACGTCCGGCGGAGAGATCCTAATCTGTTGGAACTTACCGGCC 412
Db 219 PVFGRKHSKSPALSSWGDVLLKTDVPLSSABEAECHWADTELNRRLRRRPFCSKVEGYGSV 278
QY 413 CGGAGCAATTTCTGTAACACTGGCCCGGAGGCGGTGTGCAAGGTGATTTTCCGGGTGGA 472
Db 279 CSCKDPTIEFSPDPLPDKNVLPVAVIAGNRPVLYRMLRSLLSAQGVSPQMIT---V 335
QY 473 AAGGGTACTTTGCGCGTTCCGGCGTGGTGGGCACATGCAAGGAATGCCCTTAGTGGCAG 532
Db 336 FIDGYEPEMDVVALFGLRGIOHTPIISKARVSQHYKASLTATNLFPEAKEAVVLEED 395
QY 533 CAGCAGCGAAGACACAGCGCATGGGTGTAATAATTTGTTACTGTCACCCCGCAGGTCGTG 592
Db 396 LDIADVDFFSLSIHLLEEDSLYICISAWNDQGYEHTAEDPALLYRVETMPLGLGWLRR 455
QY 593 TAGTCAGCTTAATCTTTGAAATAACGATCTAAAGGTGCGGTACGGGCGATGGCGAAA 652
Db 456 SLYKEELEPKWPTPEKLDWDMRMNPEQRGREGCLIPDVSRYHFGIVGLNNGYFHEA 515
QY 653 TTTGGAGAGCGGAACGATACCTCTGTGTGCTGGGCTAGCGGGTCACTTCTTAGTATT 712
Db 516 YFKKHKFNTPGVQLRNVDLSLKEAYEVEVHRLLEAEVLDHLSKNPCEDSFLPDTEGHTY 575
QY 713 TCAAGAAATCAACTTCGACCAACCGCTTGGACCTGTGTACACATTCGTTAAACCGGAAG 772
Db 576 VAFIRMEKDDFTTWTQLAKCLHLDVLRGNHRLWRLFRKKNHFLVVGVPASPSYKVK 635


```
Db 279 CSCKDPTPIEFSPDPLPDKNLVNPVAVIAGNRPNYLYRMLRSLLSAQGVSPQMIT---V 335
QY 473 AAGGGTACTTTGGCGGCTTCGGCGCTGGCGGCACATGCAAGGAATGCCCTTAGTGCAG 532
Db 336 FIDGYEPEPMDVVALFGLRGIQHTPIISIKNARVSQHYKASLTATFNLFPFAKFAVLEED 395
QY 533 CAGCAGCGAAGCAGCAGCATGGGTGTAAATTTGTTACTGGCACCCCGCAGGTCGTG 592
Db 396 LDIAVDFFSFLSQSIHLLEDDSLYCISAWNDQGYEHTAEDPALLYRVETMPLGLWVLR 455
QY 593 TAGTCACGTTAACTTTGAAAATAACGATGATAAGGTGCCGTACGGCGCATGGCAAAA 652
Db 456 SLYKEELEPKWPTPEKLDWDMWMMRMEQRRGECIIPDVSRSYHFGIVGLNNGYFHEA 515
QY 653 TTTGGAGAGCGGAAAGTACATTCCTGTGTCTGGCGTAGCGGGTCAGTTCCTAGATT 712
Db 516 YFKKHENTVPGVQLRNVDLSLKEAYEVEVHRLLSAEVLDHKNPCEDSFLPDTGHTY 575
QY 713 TCAAGAACTCACTTCGACCAACCGCTTGGACCCCTGGGTACACATTCGGTTAAAAACCGGAAG 772
Db 576 VAFIRMEKDDFTTTWTLAKCLHLWDLVRGNHRLWRLFRKKNHFLVGVGPASPYSVKK 635
QY 773 AACGTGCGTTGTACAAAAATA 793
Db 636 PPSVTPFIFFLEPPPKKEGAPGA 656
```

RESULT 15

```
US-10-141-761-480
; Sequence 480, Application US/10141761
; Publication No. US20030148432A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 480
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-761-480
```

```
Query Match 2.6%; Score 36.6; DB 15; Length 660;
Best Local Similarity 8.5%; Pred.No. 1;
Matches 53; Conservative 193; Mismatches 372; Indels 3; Gaps 1;

QY 173 CCGGAAACGACGTCACAAAGTCATTAGCAGTGCCCAATATTCGAATAACAAAGACGAAA 232
Db 39 CQTGAVLFLIVTVIVNLIKILDTERRAISEANEDPEPEQDYDEALGRLEPPRRRSGPRV 98
QY 233 TCGAAGTGAATGATCTTGGCGGAGAGCGGTTTACGGTTGGAAGAACGACCCGCTTT 292
```

```
Db 99 LDVEVYSSRSKVYVAVDGTTVLLEDEAREOGRGHIVILNQATGHNMAKRVDFDTPSPHEDE 158
QY 293 TCAAAACCGTATTATCATGATACGGGCTGTGTATGTCTGCTCGCAGAGGGCTGG 352
Db 159 AMVLFNMPVAPGRVLICTVKDESGFHLKDTAKALLRSLGSAQGPALGWRDWTWAFVGRKGG 218
QY 353 ATGCGCTGGCGGTCCGGTACGTCCGGGCGAGGATCCTAATCTGTGGTGAACATTACCGGCC 412
Db 219 PVFGEKHSKSPALSSMGDPVLLKTDVPLSSABEAECWADTELNRNRRRFRFCSKVEGVSV 278
QY 413 CGGAGCAATTCGTAAACTGGCCCCGGAAGCGGTGTGCAAGGTGATTTCCGGGTGGA 472
Db 279 CSCKDPTPIEFSPDPLPDKNLVNPVAVIAGNRPNYLYRMLRSLLSAQGVSPQMIT---V 335
QY 473 AAGGGTACTTTGGCGGCTTCGGCGCTGGCGGCACATGCAAGGAATGCCCTTAGTGCAG 532
Db 336 FIDGYEPEPMDVVALFGLRGIQHTPIISIKNARVSQHYKASLTATFNLFPFAKFAVLEED 395
QY 533 CAGCAGCGAAGCAGCAGCATGGGTGTAAATTTGTTACTGGCACCCCGCAGGTCGTG 592
Db 396 LDIAVDFFSFLSQSIHLLEDDSLYCISAWNDQGYEHTAEDPALLYRVETMPLGLWVLR 455
QY 593 TAGTCACGTTAACTTTGAAAATAACGATGATAAGGTGCCGTACGGCGCATGGCAAAA 652
Db 456 SLYKEELEPKWPTPEKLDWDMWMMRMEQRRGECIIPDVSRSYHFGIVGLNNGYFHEA 515
QY 653 TTTGGAGAGCGGAAACGATACATTCCTGTGTCTGGCGTAGCGGGTCAGTTCCTAGATT 712
Db 516 YFKKHENTVPGVQLRNVDLSLKEAYEVEVHRLLSAEVLDHKNPCEDSFLPDTGHTY 575
QY 713 TCAAGAACTCACTTCGACCAACCGCTTGGACCCCTGGGTACACATTCGGTTAAAAACCGGAAG 772
Db 576 VAFIRMEKDDFTTTWTLAKCLHLWDLVRGNHRLWRLFRKKNHFLVGVGPASPYSVKK 635
QY 773 AACGTGCGTTGTACAAAAATA 793
Db 636 PPSVTPFIFFLEPPPKKEGAPGA 656
```

Search completed: May 30, 2005, 01:42:31

Job time : 923 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2005, 19:56:49 ; Search time 267 Seconds
(without alignments)
8696.167 Million cell updates/sec

Title: US-10-622-893A-6
Perfect score: 1419
Sequence: 1 atggggaggttcgggtgacga.....atcacatcatcatcattaa 1419

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	685	48.3	1314	2	US-09-031-059-2
2	37.8	2.7	7218	1	US-08-232-463-14
3	35.8	2.5	390416	4	US-09-949-016-16923
4	34.4	2.4	1296	4	US-09-902-540-5414
5	34.4	2.4	34953	4	US-09-902-540-1263
6	33.8	2.4	468	4	US-09-252-991A-7672
7	33.8	2.4	1032	4	US-09-252-991A-7887
8	33.8	2.4	2739	4	US-09-252-991A-7592
9	33.8	2.4	2778	4	US-09-252-991A-7958
10	33.8	2.4	147382	4	US-09-949-016-14624
11	33.6	2.4	281	4	US-09-313-294A-4144
12	33.6	2.4	4403765	3	US-09-103-840A-2
13	33.6	2.4	4411529	3	US-09-103-840A-1
14	33.4	2.4	801	4	US-09-252-991A-9143
15	33.4	2.4	1818	4	US-09-252-991A-8941
16	33.2	2.3	4181	1	US-07-670-611-1
17	33.2	2.3	4181	1	US-08-220-674-1
18	33.2	2.3	4181	1	US-08-445-186-1
19	33.2	2.3	4181	1	US-08-446-549-1
20	33.2	2.3	4181	2	US-08-446-550-1
21	33	2.3	1431	4	US-09-489-039A-5719
22	33	2.3	3931	4	US-10-101-464A-961
23	33	2.3	3939	4	US-10-101-464A-961
24	32.8	2.3	999	4	US-09-902-540-8104
25	32.8	2.3	6890	4	US-09-902-540-828
26	32.8	2.3	4403765	3	US-09-103-840A-2
27	32.8	2.3	4411529	3	US-09-103-840A-1

C 28	32.6	2.3	2230	4	US-09-620-312D-282	Sequence 282, App
C 29	32.6	2.3	60424	4	US-09-949-016-12175	Sequence 12175, A
C 30	32.4	2.3	328	4	US-09-313-294A-3938	Sequence 3938, A
C 31	32.4	2.3	3108	4	US-09-614-221A-332	Sequence 332, App
C 32	32.2	2.3	774	4	US-09-252-991A-9809	Sequence 9809, App
C 33	32.2	2.3	1200	4	US-09-252-991A-8724	Sequence 8724, App
C 34	32.2	2.3	1347	4	US-09-489-039A-4659	Sequence 4659, App
C 35	32.2	2.3	174639	4	US-09-949-016-16509	Sequence 16509, A
C 36	32	2.3	1346	4	US-09-107-532A-261	Sequence 261, App
C 37	31.8	2.2	364	4	US-09-621-976-17202	Sequence 17202, A
C 38	31.8	2.2	1047	4	US-09-902-540-1202	Sequence 4084, App
C 39	31.8	2.2	24602	4	US-09-902-540-1202	Sequence 1202, App
C 40	31.8	2.2	35574	4	US-09-949-016-11843	Sequence 11843, A
C 41	31.8	2.2	35574	4	US-09-949-016-11843	Sequence 11843, A
C 42	31.6	2.2	1187	4	US-10-000-489-105	Sequence 105, App
C 43	31.6	2.2	1205	4	US-09-976-594-457	Sequence 457, App
C 44	31.6	2.2	129899	4	US-09-949-016-14684	Sequence 14684, A
C 45	31.6	2.2	148156	4	US-09-949-016-11776	Sequence 11776, A

ALIGNMENTS

RESULT 1
US-09-031-059-2
; Sequence 2, Application US/09031059
; Patent No. 5948659
; GENERAL INFORMATION:
; APPLICANT: KATO, NOBUO
; APPLICANT: SAKAI, YASUYOSHI
; APPLICANT: TANI, YOSHIKI
; APPLICANT: FUKUYA, HIROSHI
; TITLE OF INVENTION: RECOMBINANT FRUCTOSYL AMINO ACID
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,059
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/899,172
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 0020-4253P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1314 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1311
US-09-031-059-2

Query Match 48.3%; Score 685; DB 2; Length 1314;

[illegible]

QY 408 CCGCCCGGAGCAATTTCTGTAACCTGGCCCGG 439
| | | | | | | | | | | | | | | | | | | | |
Db 27185 GCTCACGGCGCCCTTCAACGACCTGGCGCGG 27216

RESULT 6

US-09-252-991A-7672
; Sequence 7672, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7672
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7672

Query Match 2.4%; Score 33.8; DB 4; Length 468;
Best Local Similarity 55.8%; Pred. No. 0.74; Mismatches 52; Indels 0; Gaps 0;
Matches 65; Conservative 0;
QY 342 GGAGGGCCTGGATCGCTGGCGGTCCGGGTACGTCCGGGCGAGGATCCTTAATCTGGTGA 401
| | | | | | | | | | | | | | | | | | | | |
Db 251 GGATGGGTCCAGCGGTTGGCGCCTTGCCTCTTGGACGAACCGATCAGCGCGTTGA 310
| | | | | | | | | | | | | | | | | | | | |
QY 402 ACTTACCCCGCCGAGCAATTTCTGTAACCTGGCCCGGAGGCGTGTGCAAGGTGA 458
| | | | | | | | | | | | | | | | | | | | |
Db 311 CCACCACCGCCAGGAGGATTACCGCACTGTCCACCCAGAACCGGAGCGCGGTGA 367

RESULT 7

US-09-252-991A-7887/c
; Sequence 7887, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7887
; LENGTH: 1032
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7887

Query Match 2.4%; Score 33.8; DB 4; Length 1032;
Best Local Similarity 55.6%; Pred. No. 1.3; Mismatches 52; Indels 0; Gaps 0;
Matches 65; Conservative 0;
QY 342 GGAGGGCCTGGATCGCTGGCGGTCCGGGTACGTCCGGGCGAGGATCCTTAATCTGGTGA 401
| | | | | | | | | | | | | | | | | | | | |
Db 854 GGATGGGTCCAGCGGTTGGCGCCTTGCCTCTTGGACGAACCGATCAGCGCGTTGA 795
| | | | | | | | | | | | | | | | | | | | |
QY 402 ACTTACCCCGCCGAGCAATTTCTGTAACCTGGCCCGGAGGCGTGTGCAAGGTGA 458
| | | | | | | | | | | | | | | | | | | | |

Db 794 CCACCACCGCCAGGAGGATTACCGCACTGTCCACCCAGAACCGGAGCGCGGTGA 738

RESULT 8

US-09-252-991A-7592
; Sequence 7592, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7592
; LENGTH: 2739
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7592

Query Match 2.4%; Score 33.8; DB 4; Length 2739;
Best Local Similarity 55.6%; Pred. No. 2.8; Mismatches 52; Indels 0; Gaps 0;
Matches 65; Conservative 0;
QY 342 GGAGGGCCTGGATCGCTGGCGGTCCGGGTACGTCCGGGCGAGGATCCTTAATCTGGTGA 401
| | | | | | | | | | | | | | | | | | | | |
Db 2328 GGATGGGTCCAGCGGTTGGCGCCTTGCCTCTTGGACGAACCGATCAGCGCGTTGA 2387
| | | | | | | | | | | | | | | | | | | | |
QY 402 ACTTACCCCGCCGAGCAATTTCTGTAACCTGGCCCGGAGGCGTGTGCAAGGTGA 458
| | | | | | | | | | | | | | | | | | | | |
Db 2388 CCACCACCGCCAGGAGGATTACCGCACTGTCCACCCAGAACCGGAGCGCGGTGA 2444

RESULT 9

US-09-252-991A-7958/c
; Sequence 7958, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7958
; LENGTH: 2778
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7958

Query Match 2.4%; Score 33.8; DB 4; Length 2778;
Best Local Similarity 55.6%; Pred. No. 2.8; Mismatches 52; Indels 0; Gaps 0;
Matches 65; Conservative 0;
QY 342 GGAGGGCCTGGATCGCTGGCGGTCCGGGTACGTCCGGGCGAGGATCCTTAATCTGGTGA 401
| | | | | | | | | | | | | | | | | | | | |
Db 418 GGATGGGTCCAGCGGTTGGCGCCTTGCCTCTTGGACGAACCGATCAGCGCGTTGA 359
| | | | | | | | | | | | | | | | | | | | |
QY 402 ACTTACCCCGCCGAGCAATTTCTGTAACCTGGCCCGGAGGCGTGTGCAAGGTGA 458
| | | | | | | | | | | | | | | | | | | | |
Db 358 CCACCACCGCCAGGAGGATTACCGCACTGTCCACCCAGAACCGGAGCGCGGTGA 302

```
RESULT 10
US-09-949-016-14624
; Sequence 14624, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 14624
; LENGTH: 147382
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(147382)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14624

Query Match      2.4%; Score 33.8; DB 4; Length 147382;
Best Local Similarity 52.5%; Pred. No. 53;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 406 ACCGCCCGGAGCAATTCGTAATCGCCCGGAGCGGTGTCGCAAGTGATTTTCG 465
Db 47933 ACCTGGCAGATGAATTTGGTAGATGAATTTGGTAGGCTGTTGCAAGTGAATGAA 47992

QY 466 GGTGGAAAGGTTACTTTGGCGTTCGCGCTGGCGGTCGCGGACATGCAAGATGCCCTTA 525
Db 47993 TGAATGAAAGTGAGCAAAATGGGATCAGGTGTGGCCATGACTTATCTGGAGATGTTG 48052

QY 526 GTGGCAGCAGCACCGGAAGCA 546
Db 48053 GCTGCAGAGAGGAGAGGA 48073

RESULT 11
US-09-313-294A-4144/c
; Sequence 4144, Application US/09313294A
; Patent No. 8476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4144
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700347784H1
US-09-313-294A-4144

Query Match      2.4%; Score 33.6; DB 4; Length 281;
Best Local Similarity 53.9%; Pred. No. 0.6;
Matches 69; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
```

```
QY 14 GTGACGATGATGACCTGGCTCTCGCGCTCACTAAGTCATCATCTCTCTCTGATCGTTGCTG 73
Db 246 GTGAACGACGAGATGCTCTGCGGATGGCATAGGGTCATCACCTTCCTCGCTGTCTGGAG 187

QY 74 CCGGGACTTGGGGCACCTCAACGGCTCTGCACCTCGCGCGCGCGGATATACCAACGTTA 133
Db 186 CAGTGGCTTCAGCGTCGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 127

QY 134 CCGTGTCTG 141
Db 126 TCGTTCGTG 119

RESULT 12
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match      2.4%; Score 33.6; DB 3; Length 4403765;
Best Local Similarity 59.4%; Pred. No. 2.9e+02;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 95 CGGCTCTGCACCTCGCGCGCGCGGATATACCAACGTTACCGTGTGGAGCCCTTATCTGT 154
Db 835772 CCGAGCAGCAGCCCGCGCGCGCGTACCGGACACCCCGCGGCGGCGGCGGCGGCGG 835713

QY 155 TCCTAGCGCCATCTCGCGCGGAAACGACGTGAACA 190
Db 835712 CCCTTACCGCGGTACCGCGCGGTTCGCGGATCAGCCCA 835677

RESULT 13
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
```

US-09-103-840A-1

Query Match 2.4%; Score 33.6; DB 3; Length 441529;
Best Local Similarity 59.4%; Pred. No. 2.9e+02;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 95 CGGCTCTGCACCTCGCGCGCGCGGATATACCAAGTTACCGGTGGACCCCTATCCTG 154
Db 833594 CCGAGCAGACGCCCGCGCGCGCGGTACCAACGACACCCCGCGGACCCCAAGCCG 833535
QY 155 TCCTAGCGCATCTCGCGCGGAAACGAGCTGAACA 190
Db 833534 CCCTTACCGCGTTACCGCGGTTGCCGATCAGCCCA 833499

RESULT 14

US-09-252-991A-9143/c
; Sequence 9143, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9143
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9143

Query Match 2.4%; Score 33.4; DB 4; Length 801;
Best Local Similarity 48.2%; Pred. No. 1.5;
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 62 TGATCGTTGGTGGCGGACTTGGGGACCTCAACGGCTCTGCACCTCGCGCGCGCGGAT 121
Db 671 TGGTGATCGCGCGCGGCTTCCCGGGATCAACACCGCGCTTGAGCTGGCGGACCGCA 612
QY 122 ATACCAAGTTCAGTGGCGGACCCCTATCTCTCCCTAGCGCATCTCCGCGGAAACG 181
Db 611 TCACCGACATCGTGTGCTGGAGCGCGCCACCTCGGCTTCCGCGGTACCGGGCGCAACG 552
QY 182 ACGTGAACAAAGTCATTAGCAGTGGCCCAATATTGGAATAACAAAGACGAAATCGAAAGTGA 241
Db 551 GCGGCGAGATCATGGCGCGGATCGGCCACGACCTGGAGAGATCCGCAAGGACGTGGCG 492
QY 242 ATGAGATCTTGGCGG 256
Db 491 AGGACGCGCTGGCGG 477

RESULT 15

US-09-252-991A-8941
; Sequence 8941, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8941
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8941
Query Match 2.4%; Score 33.4; DB 4; Length 1818;
Best Local Similarity 48.2%; Pred. No. 2.8;
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 62 TGATCGTTGGTGGCGGACTTGGGGACCTCAACGGCTCTGCACCTCGCGCGCGCGGAT 121
Db 215 TGGTGATCGCGCGCGGCTTCTCCGGATCAACACCGCGCTTGAGCTGGCGGACCGCA 274
QY 122 ATACCAAGTTCATTAGCAGTGGCCCAATATTGGAATAACAAAGACGAAATCGAAAGTGA 241
Db 275 TCACCGACATCGTGTGCTGGAGCGCGCCACCTCGGCTTCCGCGGTACCGGGCGCAACG 334
QY 182 ACGTGAACAAAGTCATTAGCAGTGGCCCAATATTGGAATAACAAAGACGAAATCGAAAGTGA 241
Db 335 GCGGCGAGATCATGGCGCGCATCGGCCACGACCTGGAGAGATCCGCAAGGACGTGGCG 394
QY 242 ATGAGATCTTGGCGG 256
Db 395 AGGACGCGCTGGCGG 409
Search completed: May 29, 2005, 23:33:04
Job time : 277 secs

GenCore version 5.1.6
Copyright (c) 1995 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2005, 19:23:28 ; Search time 6433 Seconds
(without alignments)
10688.318 Million cell updates/sec

Title: US-10-622-893A-6
Perfect score: 1419
Sequence: 1 atggaggcttcgggtgacga.....atccatcatcatcattaa 1419

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	846.4	59.6	1559	8 AFU82830	US2830 Aspergillus
2	685	48.3	1314	6 AR072866	AR072866 Sequence
3	685	48.3	1314	6 E14942	E14942 Aspergillus
4	685	48.3	1314	6 E14955	E14955 Aspergillus
5	685	48.3	1314	6 E27570	E27570 Thermostabl
6	685	48.3	1314	8 ATFAOA	Y09020 A. terreus m
7	657.6	46.3	1311	8 AB180733	AB180733 Aspergill
8	354.8	25.0	1526	8 AF035700	AF035700 Aspergill
9	350.4	24.7	1338	8 AB180732	AB180732 Aspergill
10	335.2	23.6	1320	6 BD173827	BD173827 Compositi
11	335.2	23.6	1320	6 E16562	E16562 Fusarium ox
12	333.6	23.5	1320	6 BD173829	BD173829 Compositi
13	332	23.4	1320	6 BD173828	BD173828 Compositi
14	332	23.4	1320	6 BD173830	BD173830 Compositi
15	127	8.9	110000	8 CR382133	Continuation (3 of
16	62.8	4.4	1314	6 AX712067	AX712067 Sequence
17	62.8	4.4	1314	8 AB116146	AB116146 Eupenicil
18	61.2	4.3	77986	8 EX897677	EX897677 Neurospor
19	57.6	4.1	1314	6 E22308	E22308 DNA encodin

20	57.6	4.1	1333	12	AF181866	AF181866 Synthetic
21	54.8	3.9	2780	8	AF416568	AF416568 Aspergill
22	53.6	3.8	1314	6	E22309	E22309 DNA encodin
23	53.6	3.8	1314	8	FJFAOP	Y09021 P.janthinel
24	50.4	3.6	1314	6	AX712065	AX712065 Sequence
25	50.4	3.6	1314	8	AB116147	AB116147 Coniocha
26	47.8	3.4	12770	8	SPAC139	AL132667 S.pombe c
27	44.8	3.2	2000	6	AX655393	AX655393 Sequence
28	44.2	3.1	125020	9	AF429315	AF429315 Homo sapi
29	44.2	3.0	43087	9	AP005155	AP005155 Homo sapi
30	43.2	3.0	110000	8	CR382138	CR382138 Debaryomy
31	40.6	2.9	406	6	AY180665	AY180665 Influenza
32	40.6	2.9	479	14	AY180666	AY180666 Influenza
33	40.4	2.8	125020	9	AF429315	AF429315 Homo sapi
34	40.2	2.8	190232	2	AC099532	AC099532 Bos tauri
35	38.8	2.7	3731	8	ENU28333	U28333 Emeritella
36	38.4	2.7	310174	1	AE016870	AE016870 Pseudomon
37	38.2	2.7	99427	2	AP003821	AP003821 Oryza sat
38	38.2	2.7	129320	2	AC151702	AC151702 Gallus ga
39	38.2	2.7	157797	8	AP005184	AP005184 Oryza sat
40	38.2	2.7	207960	10	AC124173	AC124173 Mus muscu
41	38	2.7	1246	6	AX164174	AX164174 Sequence
42	38	2.7	151986	8	AP005840	AP005840 Oryza sat
43	38	2.7	197358	8	AP004670	AP004670 Oryza sat
44	38	2.7	300684	1	AE017227	AE017227 Mycobacte
45	37.8	2.7	7218	6	166494	166494 Sequence 14

ALIGNMENTS

RESULT 1
AFU82830

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

AFU82830 1559 bp mRNA linear PLN 15-MAY-1997
Aspergillus fumigatus fructosyl amine: oxygen oxidoreductase mRNA,
complete cds.
U82830
U82830.1 GI:1769599
Aspergillus fumigatus
Aspergillus fumigatus
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
1 (bases 1 to 1559)
Takahashi, M., Fischetrieder, M. and Monnier, V. M.
Molecular cloning and expression of amadoriase isoenzyme (fructosyl
amine:oxygen oxidoreductase, EC 1.5.3) from Aspergillus fumigatus
J. Biol. Chem. 272 (19), 12505-12507 (1997)
97284723
9139700
2 (bases 1 to 1559)
Takahashi, M., Fischetrieder, M. and Monnier, V. M.
Direct Submission
Submitted (20-DEC-1996) Pathology, Case Western Reserve University,
2085 Adelbert Rd., Cleveland, OH 44106, USA

Location/Qualifiers
1. .1559
/organism="Aspergillus fumigatus"
/mol_type="mRNA"
/db_xref="taxon:5085"
174..1490
/note="amadoriase II"
/codon_start=1
/product="fructosyl amine: oxygen oxidoreductase"
/protein_id="AAC49711.1"
/translation="MAVTKSSLLIVGAGTGTSTALHLARGYNTVTLDPYVPSA
ISAGNDVKNVSISSQYSNNKDELEVNEILAEAFNGWKNDFPKFYVHDTGLMSACS
QEGDLRGVVRFGEDPNLVELTRPEQFKLAPEGVLDGDFPGWKGKFAISGASAC
RNALVAARQAQRMGVKFTGPTQGRVVTLLIFENNDDVKVGTGKGIKRAERTLCAG
ASAGQFLDFKNQLRPTATLVHIALKEPERALYKNIPIVNIENIERGPFPEPDEBERGK
ICDEHFGYTNVQSADGTWMSIPFEKTIQPKAEATRVALLKTMPLADRPFSFARI
CWCAATANREFLIDRHPQYHSLVGGCAGSGRGFKILFSIGNLIVDAMEGKVPQKIHLE

ORIGIN		IKWNPDIANRNRWRDLGRFGGPNRVMDFDHVKEMTNVQYRDISKL"	
Query Match		59.6%; Score 846.4; DB 8; Length 1559;	
Best Local Similarity		76.6%; Pred. No. 5.5e-240;	
Matches 1036; Conservative 0; Mismatches 316; Indels 0; Gaps 0;			
Qy	21	TGATGACCTGGCTCTGCGCGTCACTAAGTCATCATCTCTCTGATCGTTGGTCCGGGAC	80
Db	161	TCAGGACACGTGATGGGTACCAAGTCATCTTCCCTTTTGTATCGTGGGGCAGGCAC	220
Qy	81	TTGGGGACCTCAACGGCTCTGCACTTCGCGCGCGCGGATAPACAAAGTTTACCGTGT	140
Db	221	CTGGGGCACATCGACTGCTCTCCACCTTGGCAGCAAGAGGATACACAAATGTCAGCGTCT	280
Qy	141	GGACCCCTATCTGTCCTAGCGCCATCTCCGCGGGAACGAGTCGACAAAGTCAATTAG	200
Db	281	AGATCCCTACCCCGTTCCTTCAGCCCATCTCGGCTGGGAATGATGTGAACAGGTCACTC	340
Qy	201	CAGTGGCCAAATATTCGAATAACAAAGACGAAATCGAAGTGAATGAGATCTTGGCGGAAG	260
Db	341	CTCGGGCCATATAGCAACAAAGAGGAAATTTAGGTCAACGAGATCTTGGCTGAAGA	400
Qy	261	GGCGTTTAAACGTTTGAAGAACGACCGCTTTTCAAACCGTATTTATCATGATACGGGCT	320
Db	401	AGCGTTCAATGGCTGGAAGAACGACCCCTTGTTCAAACCAATATATCAGCATACTGGATT	460
Qy	321	GCTGATGTCGTCGTCGAGGAGGCGCTGGATCGGCTGGGCGTCCGGTACGTCGGG	380
Db	461	GCTCATGTTCGCGCTGCTCCAGGAAGGCTTGGACCGCTTGGAGTCCGTTCAAGGCCGG	520
Qy	381	CGAGGATCCCTAACTCGTGTGAACTTACCCGCGCGAGCAATTTCTGTAACCTGGCCCCGA	440
Db	521	TGAGGACCCCAACTTGTGGAATCTGACACGCGCGAGCAATTCGCAAAATAGCTCTGGA	580
Qy	441	AGCGGTGTTCAAGGTGATTTTCCGGGTGGAAGGCTACTTTGCGGTTTCGCGCGCTGG	500
Db	581	GGGTGTTCTACAGGAGATTTCCCGGCTGGAAGGGCTACTTTGCGGTTTCAGGAGCTGG	640
Qy	501	CTGGGCATGCAAGNATGCTTAGTGGCAGCAGCAGCGGAGCAGCAGCGCATGGGTGT	560
Db	641	TTGGGGCCATGCTCGCAATGCACTCGTGGCTGCTGCAAGGGAGGCTCAGAGAAATGGCGT	700
Qy	561	AAATTTGTTACTTGGCACCCGCGAGGCTCGTGTAGTCACGTTTAACTTTGAAAATAACGA	620
Db	701	GAAGTTCGTAATCTGGCACTCTCAGGSCAGAGTAGTCATCTATATTGAGAAATACGA	760
Qy	621	TGTAAGGTGCGCTTACGGGCGATGCAAAATTTGGAGAGCGGAACGTACATTCCTGTG	680
Db	761	TGTCMAAGGTGCGGTTACCGGAGACGGCAAGATTTGGCGTGCAGAGCGCACATTCCTCTG	820
Qy	681	TGCTGGGCTAGCGCGGTGAGTTCCTAGATTTCAAGATCAACTTCGACCAACCGCTTG	740
Db	821	CGCGGTGCGAGCGGCTGGTCACTTCCTGCACTTCAAGAAATCAGTTGGCTCCACCGGATG	880
Qy	741	GACCTGGTACACATTCGCTTAAACCGGAAGACGCTGCTGTGTACAAAATATACCGGT	800
Db	881	GAGCTGGTTTCAATTTGCTCTGAGGCTGAGGAGCGGGCTCTTTACAGAAATATCCAGT	940
Qy	801	TATCTTTTAAACATCGAACCGGGGTTTTCTTTTGAACCCGATGAGAGCGCGGTGAGATTAA	860
Db	941	TATCTTCAACATTTGAGAGGGGTTCTTCTTCGAACCAAGATGAGGAGCGGCTGAGATTAA	1000
Qy	861	AATATGGGATGAACACCGGGCTACACAATATGGTCCAGATGTCAGACCGGCAGATGAT	920
Db	1001	GATCTGCGACGAACATCCGGGGTATACCAATATGGTACAGTCTGCGCAGCGGCAAGATGAT	1060
Qy	921	GAGCATTTCCGTTCGAAAAACCCAGATTTCCAAAAGAGCGAAACGCGGCTTCGGGCCCT	980
Db	1061	GAGCATTTCTTTTGAAGAAAGTCTAGATTTCTTAAAGAAAGCCGAGACGAGGGTTAGGCTCT	1120
Qy	981	GCTGAAAGAGACAAATGCCAGCTGGCAGACCGGTTCATTCAGCTTCGCAACGCAATTCGTG	1040

Db	1121	GCTTAAAGAGACGATGCCACAGCTTCGACACCGTCCATTCAGTTTCGCCACGAGATTTGCTG	1180
Qy	1041	GTGTGCCGATACCGCAATCGCGAAATTCCTGTAGATAGATCCGACATCCGACGATCCACAGTCT	1100
Db	1181	GTGCGCGGACACTGCCAACCGGAGTTCCTGATCGATCGCATCTCTAGTACCATTTGCGCT	1240
Qy	1101	TGTGTTGGGCTGTGTGTGCGAGCGGAAGAGGGTTTAAATATCTGCCCTTCTATTGCGGAATCT	1160
Db	1241	TGTGCTGGGCTGCGGGCTTCGCGCAGAGGATTCAAATATCTACCTTCAATTGGCAATCT	1300
Qy	1161	CATTGTTGACCGATGGAAGGTAAGTGGCGGCNAAAAATTCAGAAATTAATCAAGTGGAA	1220
Db	1301	CATCGTTGATGTATGGAAGGCAAGGTCCTCAAAGATCCACGAATGATTTAAATGGAA	1360
Qy	1221	CCCGACATTTGGCGCAACCGTAACCTGGCGTGATCTCTGGGCGGTTTGGCGGTCCTCAA	1280
Db	1361	CCAGATATTGTCGCCATTCGCACTGGAGGATCTTTGGGGAGATTCGGGGGTCCAA	1420
Qy	1281	TGCTGTGATGGAATTTTCATGATGTGAAGGAATGGACCAATGTTTCAGTATCGTGATTTTC	1340
Db	1421	CAGAGTAATGGACTTCCACGAGCTCAAGGAGTGGACAAATGTACAAATATAGAGATATTTTC	1480
Qy	1341	CRAAGTGAAGAGGAGTTCGAAGGTAAGCCAA	1372
Db	1481	CRAAGTTAATGATGATGACAAATCTGCCCAA	1512

RESULT 2

AR072866

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

ORIGIN

Sequence 2 from patent US 5948659.

AR072866

AR072866.1 GI:9999629

Unknown.

Unknown.

1 (bases 1 to 1314)

Kato, N., Sakai, Y., Tani, Y. and Fukuya, H.

Recombinant fructosyl amino acid oxidase

Patent: US 5948659-A 2 07-SEP-1999;

Location/Qualifiers

1..1314

source

/organism="unknown"

/mol_type="unassigned DNA"

Query Match

Best Local Similarity

Matches

48.3%; Score 685; DB 6; Length 1314;

70.7%; Pred. No. 5.7e-192;

926; Conservative 0; Mismatches 380; Indels 3; Gaps 1;

305 CACAGGAAGGATTGGAGCGTCTGGGGTCCGCGTGCAGCCTGAAGATGAACCCGATGTAG 364
Qy TGGAACTTACCGCCCGGAGCAATTTCTGTAATCTGGCCCGGAGAGGGCTGTGCAAGGTG 457
Db CCGAAATTTGACTCGCGCGGAGCAGTTCGCCAGCTGGCCCCC---GGCGTCTTGAAGGTA 421
Qy ATTTTCCGGGTGCGAAAGGTACTTTTCGGGTTCCGGCTCGGCTGGGCAATGCAAGGA 517
Db ACTTCCCGGTGAGAGGGGTACACATTTCTGCTCAAGCGGGCTGGCGCATGCGCGCA 481
Qy ATGCTTTAGTGGCAGCAGCAGCGAAGCAGCAGCGCATGGGTGTAATAATTTTACTTGSCA 577
Db ACGCCTTGGTCCGCGCGCGGAGGAGCAGCAGCGCTGGGTGCTGCTGCGCGGAT 541
Qy CCGCGAGGCTGTGTAGTACAGTTATCTTTGAAATACAGATGTAAGAGGTGCGCGTTA 637
Db CGCGCAGGCGCAGATCATCAGTTGATTTTGGAGAAACAAGCATGTGAAGGGTGGCGTCA 601
Qy CGGCGATGCAAAATTTGAGAGCGCAACGTACATTTCTGTGTGCTGGGCTAGCGCG 697
Db CGGCGAGCGAAGATCTGGCGGCGGAGCAGACTATCTCTGCGCTGTGGCGCGCG 661
Qy GTCAGTTCTTAGATTTCAAGAAATCACTTCGACCAACCGCTTGGACCTTGGTACACATTG 757
Db GCCAGTTTCTGGATTTCAAGAGCAACCTGCGTCCCACTGCTGGACTCTGTCACATCC 721
Qy CGTTAAACCGGAGAACGCTGCTGTGTAACAAATATACCGGTTATCTTTAAACATGAAC 817
Db AGTTGAAGCGGAGAGCGTCCCAAGTATAAACAATCAACATGCGGCTGTTCAACATCGA 781
Qy GGGGTTTTTCTTTGAACCGATGAGGAGCGGCTGAGATTAAATATGCGATGACAC 877
Db AGGGTTCTTCTCGACCGGATGAGGAGGCTGTGTAATCAAGATCTGCGACCAACC 841
Qy CGGCTACACAAATATGCTCCAGAGTCAGACGCGCATCATGATGAGCATTCCTGTTGAAA 937
Db CCGGTACAGATATGACCAACGGGGCCGACGCGCGCTGAGGAGCATTCCTTTCGAGA 901
Qy AATCCAGATTCCAAAGAGCGAAACGGCGGTTTGGGCGCTTCTGCTGAAAGAGACATGC 997
Db AGACGCGAGTTCTCTCGAAGAGCGAGATGCGCGTCCGCAAGCTTCTGCTGAAACGATGC 961
Qy CCAGCTGGCAGACCGTCCATCAGCTTCGACGATTCGCGATTCGCTGTTGGGCTGTGGTG 1057
Db CTAGCTTGGGACCGCGCGCTTCAAGTTTCGAAAGGATCTGCTGTGTGCGGATACCCCA 1021
Qy ATCGCAATTTCTGATAGATCGACATCCGAGTACACACAGTCTTGTGTTGGGCTGTGGTG 1117
Db ATCGAGTTTATCATTTGACCGTCAATCCGATACCGCTGCTGTTGTTGGTGTGGTG 1081
Qy CGAGCGAAGAGGTTTAAATATCTGCTTCTATTTGGGAATCTCATTTGACCGCATGG 1177
Db CTTTCAGGACGAGGCTTCAAAATATCTTCCCTCGATCGGAAGCATCATCGCAGACGCCATGG 1141
Qy AAGTAAAGTGGCGCAAAATTTCAAGATTAATCAAGTGGAAACCGGACATTTGGCGGA 1237
Db AGGACAAACCCCGGCAAAATTTCAAGTGTATCGCTGAGCGCGGAAATCGCGATCA 1201
Qy ACCGTAACTGGCGTGAATCTCTGGGGGTTTGGCGGTCAAAATCTGTTGATGATTTTC 1297
Db ACCGTAACTGGGGGACAGATTAGTTCATTTGGAGGGCCCAACCGGCTCATGATTTCA 1261
Qy ATGATGTGAAGGAATGACCAATTTTCAGTATCTGATATTTTCCAGCT 1346
Db ATGAAGTGAAGGATGGACTTAATGTACCCCAAGGAGACATCTCGAAGTT 1310

RESULT 3

E14942
LOCUS Aspergillus terreus mRNA for fructosyl amino acid oxydase (FAOD-L).
DEFINITION E14942
ACCESSION E14942
VERSION E14942.1 GI:5709625

KEYWORDS JP 1998033177-A/1.
SOURCE Aspergillus terreus
ORGANISM Aspergillus terreus
REFERENCE 1 (bases 1 to 1314)
AUTHORS Kato, N., Sakai, Y., Tani, Y. and Fukuya, H.
TITLE FRUCTOSYL AMINO ACID OXIDASE
JOURNAL Patent: JP 1998033177-A 1 10-FEB-1998;
KDK CORP
COMMENT OS Aspergillus terreus
PN JP 1998033177-A/1
PD 10-FEB-1998
PF 23-JUL-1996 JP 1996193344
PI KATO NOBUO, SAKAI YASUYOSHI, TANI YOSHIKI, FUKUYA HIROSHI PC
C12N15/09, C07H21/04, C12N1/21, C12N9/06, C12P21/02, C12Q1/26, PC
(C12N1/21,
PC C12R1:19), (C12N9/06, C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1314
FT /strain= 'GPI' /clone= 'pFAL2'
FT CDS 1..1314
FT /product= 'fructosyl amino acid oxydase'.
FEATURES
source 1..1314
/organism= 'Aspergillus terreus'
/mol_type= 'genomic DNA'
/db_xref= 'taxon:33178'
ORIGIN
Query Match 48.3%; Score 685; DB 6; Length 1314;
Best Local Similarity 70.7%; Pred. No. 5.7e-192;
Matches 926; Conservative 0; Mismatches 380; Indels 3; Gaps 1;
Qy 38 CCCTCACTAAGTCAATCTCTCTGATCGTGGTCCGGGACCTTGGGGCACCTCAACGG 97
Db 5 CAGTCAACCAAGTCTCTGTCGATATTGATCATCGGGCGGGCACCTGGGGTGTCTCAACTG 64
Qy 98 CTCTGACCTCGCGCGCGCGGATATACCAAGTTACCGTGTGGACCCCTATCTGTCC 157
Db 65 CCCTGATCTTGGCGCAGAGGATACCAATGTCATCTGTCCTGACCCGCTACCGGTTTC 124
Qy 158 CTAGCGCATCTCGCGGAAACGACGTGAACAAAGTCATTAGCAGTGGCCAAATATTCGA 217
Db 125 CATCAGCCATTTCGGCGGCAACGACGTCAACAGATCATCTGTCGCGCCAGTACAGCA 184
Qy 218 ATAAACAAAGACGAAATCGAAGTGAATGAGATCTTGGCGGAAGAGCGGCTTTAAACGGTTGA 277
Db 185 GCAAGAAGGACGAGGTGCAAGTCAATGAGATATCGCCGAACAGGCCCTTCAATGCGTGA 244
Qy 278 AGAACGACCGCTTTTCAACCGTATATCATGATACGGCGCTGCTGATGCTGCTGTCT 337
Db 245 AAAATGACCCCATCTTCAAGCCGTACTACCAAGACCGCGCTGCTGATGTCGCGCACCA 304
Qy 338 CGCAGAGGCGCTGATCGCTCGGGGTTCGGGTTCGCGGTGCGACCTGAAGATGAACCCGATGTAG 397
Db 305 CACAGGAAGGATTTGGAGCGTCTGGGGTTCGGGTTCGCGGTGCGACCTGAAGATGAACCCGATGTAG 364
Qy 398 TGAACCTTACCCCGCGGAGCAATTTCTGTAATCTGGCCCGGAGCGGTGTGCAAGGTG 457
Db 365 CCGAATTTGACTCGCGCGGAGCAGTTCGCGCAGCTGGGCCCC---GGCGTCTTGAAGGTA 421
Qy 458 ATTTTCCGGGTTGGAAGGGAATTTTGGCGGCTTCGCGCGCTGCTGGGCAATGCAAGGA 517
Db 422 ACTTCCCGGTTGGAGGGGTACACATTCGCTCAAAACGCGGGCTGGCGCATGCGGCA 481
Qy 518 ATGCTTAGTGGCAGCAGCAGCGGATGCGGATGCGGCTGGAATTTTGTACTGCA 577

Db	482	ACGCCCTGTCGCGCGCGCGGAGGCACAGCCCTGGGTGTGCGCTTCGTTCGCGGGAT	541
Qy	578	CCCGCAGGTCGTGTAGTCACGCTTAATCTTTGAAAATAACGATGTAAAGGTCGCGTTA	637
Db	542	CGCGCAGGCGAGTCATCAGTGTGATTTTGGAGAACACGATGTGAGGGTTCGCGTCA	601
Qy	638	CGGCGATGCGCAAAATTTGGAGAGCGGAACGTAATCTCTGTGTCTGGGGCTAGCGCG	697
Db	602	CGCGGACGCGCAAGATCTGCGCGCGCGAGCAGACTATCTCTCGCTGGTGGCGCGCG	661
Qy	698	GTCACTCTCTAGATTTTCAGAAATCAACTTCGACCAACCGCTTGGACCTGGTACACATG	757
Db	662	GCAGTCTCTGGATTTCAAGGACCAACTCGCTCCCACTCGCTGGACTTGGTCCACATCC	721
Qy	758	CGTTAAACCGGAAGACGTCGTGTACAAAATATACCGGTATCTTTAAACATCGAAC	817
Db	722	AGTTGAAGCGGAAGAGCTGCCAGTATANAACHATGCCGGTGTCTTCAACATCGAGA	781
Qy	818	GGGGGTTTTCTTTGAACCCGATGAGGAGCGCGTGAGATTAANAATATCGGATGAACCC	877
Db	782	AGGGGTTCTTTCGAGCCGGATGAGGAGCGTGGTGAATCAAGATCTGCGACGAACACC	841
Qy	878	CGGGCTACACAAATATGTTCCAGATGTCAGACCGCAGGATGATGAGATTCGTTGAAA	937
Db	842	CCGGGTACACGAATATGACACCGGGGCGACGCGCGGTGAGGAGCATTTCCCTTCGGA	901
Qy	938	AAACCCAGATTTCCAAAGAGCGGAACGCGCTTCGGGCCCTGCTGAAGAGACAAATGC	997
Db	902	AGACGAGGTTCTTCGAGAAAGCGAGATGCGCGCTCCGCAAGCTTCTGTCTGAACGATGC	961
Qy	998	CCAGCTGGCAGACCGTCCAATCAGCTTCGACGCAATTTGCTGGTGTGCCGATACCGCGA	1057
Db	962	CTCAGCTTGGGACCGCGCGTTCAGTTTCGCAAGGATCTGCTGTGTGGGATACCCCA	1021
Qy	1058	ATCGCAATTCCTGATGATGACATCCGCGAGTACACAGTCTTGTGTGGGCTGTGGTG	1117
Db	1022	ATCGCAGTTTATCAITGACCGTCATCCGGAATACCGCTGCTGTGTTGGGTGTGGTG	1081
Qy	1118	CGAGCGGAGAGGTTTAAATATCTGCTTCTATTGGGAATCTCATTTGTGACCGATGG	1177
Db	1082	CTTCAGACGAGGCTTCAAAATATCTTCCCTCGATCGGAAGCATCATCGAGCGCATGG	1141
Qy	1178	AAGTAAAGTCCGCGCAAAAATTCAGGAATTAATCAAGTGAACCCCGGACATTCGCGCGA	1237
Db	1142	AGGACAAAACCCCGCAAAAATCCACAAGTGTATCCGCTGGAGCCCGGAATCGCGATCA	1201
Qy	1238	ACGTAATCGGCTGATCTCTGGGGGTTTGGCGGTCCAAATCTGTGATGATGATTTTC	1297
Db	1202	ACCGTAATCTGGGGGACAGATTAGTTCGATTTGGAGGGCCCAACCGGGTCAATGGATTCA	1261
Qy	1298	ATCATGTGAAGGAATGACCAATGTTTCAGTATCGTGATATTTCCAAAGCT	1346
Db	1262	ATGAAGTGAAGGAGTGAGCTAATGTACCCCAAGGGACATCTCGAAGTT	1310
RESULT 4			
LOCUS	E14955	1314 bp	DNA linear PAT 28-JUL-1999
DEFINITION	Aspergillus terreus mRNA for fructosyl amino acid oxydase (FAOD-L).		
ACCESSION	E14955		
VERSION	E14955.1	GI:5709638	
KEYWORDS	JP 1998033180-A/1.		
SOURCE	Aspergillus terreus		
ORGANISM	Aspergillus terreus		
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;		
AUTHORS	1 (bases 1 to 1314)		
TITLE	RECOMBINED FRUCTOSYL AMINO ACID OXIDASE		
JOURNAL	Patent: JP 1998033180-A 1 10-FEB-1998;		
COMMENT	KDK CORP		
	OS Aspergillus terreus		
	PN JP 1998033180-A/1		

PD	10-FEB-1998	24-JUL-1996	JP 1996194557	
PI	KATO NOBUO, SAKAI YASUYOSHI, TANI YOSHIKI, FUKUYA HIROSHI	PC	C12N15/09, C07H21/04, C12N1/19, C12N9/06, C12P21/02, C12Q1/26, PC	
G01N33/66,				
PC	G01N33/68, (C12N1/19, C12R1:72), (C12N9/06, C12R1:72); CC			
strandedness:	Double;			
CC topology:	Linear;			
FH	Key	Location/Qualifiers		
FT	source	1..1314	/organism='Aspergillus terreus' FT	
FT	/strain='Gp1'	/clones='pFAL2'		
FT	CDS	1..1314	/product='fructosyl amino acid oxydase'.	
FT	Location/Qualifiers			
FEATURES				
source				
1..1314				
/organism="Aspergillus terreus"				
/mol_type="genomic DNA"				
/db_xref="taxon:33178"				
ORIGIN				
Query Match 48.3%; Score 685; DB 6; Length 1314;				
Best Local Similarity 70.7%; Pred. No. 5.7e-192;				
Matches 926; Conservative 0; Mismatches 380; Indels 3; Gaps 1;				
Qy	38	CGTCACTAAGTATCATCTCTCCTGATCGTTGGTCCGGGACTTGGGCGACCTCAACGG	97	
Db	5	CAGTCAACCAAGTCTTCGTGATATTGATCATCGGGCGGGCACCTGGGGTGTCTCAACTG	64	
Qy	98	CTCTGACACTCGCGCCCGGATATACCAACGTTACCGTGTGGACCCCTATCTGTCTTC	157	
Db	65	CCCTGATCTTGGCGGAGAGGATACCAATGTCACTGTCTTGAACCCGCTACCCGCTTC	124	
Qy	158	CTAGCGCCATCTCCGCGGAAACGACGTGAAACAAAGTCAATAGCATGTGGCAATATTCGA	217	
Db	125	CATCAGCCATTTCCGCGCGCAACGCTCAACAAAGATCATCTCGTCCGCGCAGTACAGCA	184	
Qy	218	ATACCAACGACCAATCGAAGTGAATGAGATCTTGGCGGAAGAGGCTTTAACGGTTGGA	277	
Db	185	GCAAGAAGGACGAGGTGGAAGTCAATGAGATATCCGCCAAGAGGCTTCAATGGCTGGA	244	
Qy	278	AGAACGACCCGCTTTTCAAAACCGTATATCATGATACGGGCTCTGTGCTGTGCT	337	
Db	245	AAAATGACCCCATCTTCAAGCCGTACTACACGACACCGGCGTGTGATGTCCGCCACCA	304	
Qy	338	CGCAGAGGCGCTTGATTCGCTGGGCTCGGGGTAAGTCCGGGCGAGGATCCTAATCTGG	397	
Db	305	CACAGGAAGGATGAGAGCGTCTGGGGGTCCGCGTGGGACCTGAAGATGAACCCGCGATAG	364	
Qy	398	TGGAACTTACCCCGCGGAGCAATTTCTGTAACCTGGCCCGGAGCGGTGTTCAAGGTG	457	
Db	365	CGAATTGATCGGCGGAGCAGTTCGCCAGCTGGCCCCC-----GGCGTCTTGAAGGGTA	421	
Qy	458	ATTTTCGGGTGGAAGGGTACTTTTGGCGGTTCGGCGCTGGCTGGGCGACATCGCAAGGA	517	
Db	422	ACTTCCCGGTTGGAGGGGTACCATTTGCTCAAAACGGGGCTGGCGCATCGCGCA	481	
Qy	518	ATGCTTTAGTGGCAGCAGCAACGCGAAGCAACGCGCATGGGTGTAATTTGTTACTGGCA	577	
Db	482	ACGCCCTGGTCCGCGCGCGGAGGCAACGCGCTGGGTGTGCGCTTCGTCGCGGGAT	541	
Qy	578	CCCGCAGGGTCTGTAGTACGTTAATCTTTCAAAATACGATGTAAAGAGGTGCCGTTA	637	
Db	542	CGCGCAGGGCAGAGTCATCACGTTGATTTTGAAGAACACGATGAAGGGTGCCGTC	601	
Qy	638	CGGCGCATGGCAAAATTTGGAGAGCGGAACGTAATTTCTGTGTGTGGGGCTAGCGCGG	697	
Db	602	CGCGGACGCGAAGATCTGGCGGCGGAGCAGACTATCTCTCGCTGGTGGCGCGCCG	661	
Qy	698	GTCACTTCTTAGATTTCAAGAATCAACTTTCGACCAACCGCTTGGACCCCTGGTACACATG	757	

Db 662 GCCAGTTTCTGGATTTCAGGACCAACCTGCTCCCACTGCTGGACTCTGTCCACATCC 721
Qy 758 GGTAAACCCGGAAGAGTGGTGTATACAAATATACCGTGTATCTTTAAATCGAAC 817
Db 722 AGTTGAAGCCGGAAGAGCGTCCCACTATATAAATCATCCCGTGTCTTCAACATCGAGA 781
Qy 818 GGGGTTTCTTTCTTTGAACCCGATGAGGAGCGGCTGAGATTAAATATGCGATGAACACC 877
Db 782 AGGGGTTCTTCTCGAGCCGATGAGGAGCGTGTGAATCAAGATCTGCGACGAACACC 841
Qy 878 CGGCTACACAAATATGTTCCAGAGTCAGACGCGGATGATGAGCATTCCTGTTCGAAA 937
Db 842 CCGGTTACACAAATATGACCAACGCGGCGGAGCGCTGAGGAGCATTCCTTTCGAGA 901
Qy 938 AAACCCAGATTCCTTGAACCCGATGAGGAGCGGCTGAGATTAAATATGCGATGAACACC 977
Db 902 AGACGCGAGTTCTTCTGAAGAGCGGAGATGCGCGTCCGCAAGCTTCTGTCTGAACGATGC 961
Qy 998 CCCAGCTGGCAGACCGTCCATTCAGCTTCGCAACGATTTGCTGTGTCGCGATACCGGA 1057
Db 962 CTGAGCTTGGGACCGGCGGTTGAGTTTCGCAAGATCTGCTGTGTCGCGATACCGCCA 1021
Qy 1058 ATCCGGAATCTCTGATGATCGACATCCGAGTACCAAGTCTTGTGTTGGGCTGTGGTG 1117
Db 1022 ATCCGAGTTTATCATTCAGCGTCTATCCGAATACCGTCTGCTTGTCTTGGGTTGGTG 1081
Qy 1118 CGACCGAAGAGGTTTAAATATCTGCTTCTTATTTGGGAATCTCATTTGTCACCGGATGG 1177
Db 1082 CTTGAGGACGAGGCTTCAAAATATCTTCCCTGATCGGAAGCATCATCGCAGACCCATGG 1141
Qy 1178 AAGGTAAAGTCCGCAAAATTCACGAATTAATCAAGTGAACCCGAGATTCGGGCGA 1237
Db 1142 AGGACAAACCCCGGCAAAATTCACGAATTAATCAAGTGAACCCGAGATTCGGGCGA 1201
Qy 1238 ACCGTAACTGCGGTGATCTCTGGGGGTTTGGCGGTCCAAATCTGTGATGATTTTC 1297
Db 1202 ACCGTAACTGCGGTGATCTCTGGGGGTTTGGCGGTCCAAATCTGTGATGATTTTC 1261
Qy 1298 ATGATGTGAAGGAATGACCAATGTTGATATCGTATGATTTTCAAGCT 1346
Db 1262 ATGAAGTGAAGGAGTGAGTAAATGTCACCCAAAGGAGCATCTCGAAGTT 1310

RESULT 5

E27570 LOCUS 1314 bp DNA linear PAT 18-JUN-2001
DEFINITION Thermotable fructosylamino-acid oxidase.
ACCESSION E27570
VERSION E27570.1 GI:13018233
KEYWORDS JP 1999221081-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1314)
AUTHORS Yoshiki,T., Toru,K., Nobuyuki,Y., Akio,K. and Hiroshi,F.
TITLE Thermotable fructosylamino-acid oxidase
JOURNAL Patent: JP 1999221081-A 1 17-AUG-1999;
KYOTO DAIICHI KAGAKU CO LTD
COMMENT OS Aspergillus terreus Gp1
PN JP 1999221081-A/1
PD 17-AUG-1999
PF 10-FEB-1998 JP 1998028169
PR
PI YOSHIKI TANI,TORU KATSURAGI,NOBUYUKI YOSHIDA,AKIO KATO, PI
HIROSHI FUKUIE
PC C12N15/09,C07K19/00,C12N1/21,C12N9/06/(C12N15/09,C12R1:66),
PC (C12N1/21,C12R1:19),(C12N9/06,C12R1:19),C12N15/00,(C12N15/00,
PC C12R1:66)
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..1314
FT /organism='Aspergillus terreus Gp1'.

FEATURES

source
1..1314
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 48.3%; Score 685; DB 6; Length 1314;
Best Local Similarity 70.7%; Pred. No. 5.7e-192;
Matches 926; Conservative 0; Mismatches 380; Indels 3; Gaps 1;
Qy 38 CCGTCACCTAGTCATCATCTCTCGTATCGTGGCGGACCTTGGGGCACCTCAACGG 97
Db 5 CAGTCACCAAGTCTTCTGATATGATCATCGGGCGGGCACCTTGGGGTGTCTCACTG 64
Qy 98 CTCTGCACCTCGCGCGCGGAGATATACCAAGTGTACCGTGTGACGCCCTATCTGTCC 157
Db 65 CCCTGCATCTTGGCGCGAGAGGATACACCAATGTCACTGTCTTGGACCGTACCGGTT 124
Qy 158 CTAGCGCATCTCGCGCGGAAACGACGTGAACAAAGTCAATTAGCAGTGGCCAAATTCGA 217
Db 125 CATCAGCCATTTCCGCGCGCAACGACGTCAACAAAGATCATCTCGTCCGGCCAGTACA 184
Qy 218 ATAAACAAGACGAATCGAAGTGAATGAGATCTTTGGCGGAAGAGCGTTTAAACGTTG 277
Db 185 GCAAGAAGACGAGTCTGAAGTCAATGAGATTAATCGCGCAACAGGCTTCAATGCTGGA 244
Qy 278 AGAACGACCCGCTTTTCAAAACCGTATTAATCATGATACGGGCTCTGATGTCTGTCTG 337
Db 245 AAATGACCCCATCTTCAAGCGGTACTACCAACGACCGCGCTGATGTCCGCCACCA 304
Qy 338 CGCAGGAGGCTGATCGCTGGCGTCTCGGGTACGTCCGGGAGAGATCTTAATCTGG 397
Db 305 CACAGGAAGATTGGAGCGTCTGGGGTCCGCGTGCACCTGGAAGATGAACCCGATGTAG 364
Qy 398 TGGAACCTTACCGCGCGGAGCAATTCGTAACCTGCGCCCGGAAGAGGTGTTCAAGGTG 457
Db 365 CGGAATTGATCGCGCGGAGCAGTTCCGCAAGCTTGGCCCCC ---GGGCTTTGAAGGTA 421
Qy 458 ATTTTCCGGTGGAAAGGTTACTTTTCCGGCTTCCGGCTGGTGGGCAATGCAAGGA 517
Db 422 ACTTCCCGGTTGGAGGGGTACCACATTCGCTCAACCGGGCTGGCGCATGGCGCA 481
Qy 518 ATGCTTAGTGGCAGCAGCAGCGAAGCAGCCATGGGTGTAATTTGTTACTGCA 577
Db 482 ACGCCCTGTGTCGCGCGCGCGGAGGACAGCGCTGGGTGTCGCTTCGTCGGGAT 541
Qy 578 CCGCGCAGGCTGTGTAGTCACGTTAATCTTTGAATAACGATGAAGGTCCCGTTA 637
Db 542 CGCGCAGGGCAGAGTCATCACGTTGATTTTGAAGAACACGATGTGAAGGGTCCCGTCA 601
Qy 638 CGGCGCATGGCAAAATTTGGAGAGCGGAACGTATCTCTGTGCTGGGGCTAGCGCG 697
Db 602 CGCGCAGCGCAAGATCTGGCGGCGCAGCAGATATCTTCTGCGCTGGTGGCGCGCG 661
Qy 698 GTGAGTCTTAGATTTCAAGATCAACTTCGACCAACCGCTTGGACCGCTGTGACATTTG 757
Db 662 GCCAGTTTCTGGATTCAAGGACCACTGCGTCCCACTGGGTGAGTCTGGTCCACATCC 721
Qy 758 CGTTAAACCGGAAGACGTGCTGTGTACAAAATATACCGGTTATCTTTAAACATCGAAC 817
Db 722 AGTTGAAGCCGGAAGAGCGTCCCAAGTATATAAACAATGCGCGGTGGTCTTCAACATCG 781
Qy 818 GGGGTTTTTTCTTTTGAACCCGATGAGGAGCGGCTGAGATTAAATATGCGATGAACACC 877
Db 782 AGGGTTCTTCTTCGAGCCGATGAGGAGCGTGGTGAATCAAGATCTGCGACGAACACC 841
Qy 878 CGGGCTACACAAATATGTTCCAGAGTCAGACGCGCAGATGATGAGCATTCCTGTTGAAA 937
Db 842 CCGGTTACACGAATATGACCAACGCGGGCGGAGCGCTGAGGAGCATTCCTTTCGAGA 901
Qy 938 AAACCCAGATTCCTTGAAGAGCGGAACGCGCTTGGGGCCCTGCTGAAGAGACATGC 997

```

Db      902 AGACGAGGTTCTCTCGAGAGCGGAGATCGCGTCCGCAAGCTTCTGTCTGTAACGATCG 961
QY      998 CCAGCTGGCAGACCGTCCATTCCAGCTTCGCGACGATTTGCTGTGCGGATACCGCGCA 1057
Db      962 CTAGCTTGGGACCGCGCTTCAGTTTTCGAGGATCTGTGTGCGGATACCGCCCA 1021
QY      1058 ATCGCGAATTTCTGTAGATCGACATCCGCGAGTACCACAGTCTTGTGTGGGCTGTGGTG 1117
Db      1022 ATCGCGATTATCATTTGACCGTTCATCCGATACCGTTCGTTGTTCTGGTGTGGTG 1081
QY      1118 CGAGCGAGAGGGTTAAATATCTGCTTCTATTCGGAATCTCATTTGTGAGCGCATGG 1177
Db      1082 CTTTCAGGACGAGGCTTCAATATCTTCCCTCGATCGGAGGATCATCGCAGACGCCATGG 1141
QY      1178 AGGTAAAGTGGCGCAAAATTCAGCAATTAATCAAGTGGACCGGACATTCGCGCGCA 1237
Db      1142 AGGACAAAACCCCGGCAAAATCCAAAGCTGATCCGCTGGAGCCCGGAAATCGCGATCA 1201
QY      1238 ACCGTAACCTGGCGTGATCTCTGGGCGTTCGCGGTCGCAATCGTGTGATGGATTTTC 1297
Db      1202 ACCGTAACCTGGGCGGACAGTTAGTTCGATTTGAGGCGCCCAACCGGTCTGATTTCA 1261
QY      1298 ATGATGTGAAGAAATGGACAATGTTTCAGTATCGTGATATTTCCAAAGCT 1346
Db      1262 ATGAAGTGAAGAGTGGACTAATGTCAACCAAGGACATCTCGAAGTT 1310

```

RESULT 6

```

ATFAOA      1314 bp mRNA linear PLN 16-APR-1997
LOCUS      A.terreus mRNA for fructosyl amino acid oxidase.
DEFINITION Y09020
ACCESSION Y09020.1 GI:1653969
VERSION faoA gene; Flavoprotein; fructosyl amino acid oxidase.
KEYWORDS Aspergillus terreus
SOURCE Aspergillus terreus
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
          Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1 Yoshida,N., Sakai,Y., Isogai,A., Fukuya,H., Yagi,M., Tani,Y. and
          Kato,N.
          Primary structures of fungal fructosyl amino acid oxidases and
          their application to the measurement of glycosylated proteins
          Eur. J. Biochem. 242 (3), 499-505 (1996)
JOURNAL 97175014
MEDLINE 9022674
PUBMED
REFERENCE 2 (bases 1 to 1314)
          Kato,N.
          Direct Submission
          Submitted (23-OCT-1996) N. Kato, Fac. of Agriculture, Kyoto Univ.,
          Kitashirakawa-oiwake-cho, Sakyo-ku, Kyoto 606-01, JAPAN
JOURNAL Location/Qualifiers
          1. .1314
             /organism="Aspergillus terreus"
             /mol_type="mRNA"
             /isolate="GPI"
             /db_xref="taxon:33178"
             /cell_line="E.coli SOLR"
             1. .1314
             /gene="faoA"
             1. .1314
             /gene="faoA"
             /codon_start=1
             /product="fructosyl amino acid oxidase"
             /protein_id="CAA70218.1"
             /db_xref="GI:1653970"
             /db_xref="GOA:Q92204"
             /db_xref="UniProt/TsEMBL:Q92204"
             /translation="MPVTKSSSILLIGAGTWGCGSTALHLARCYTNVTLDPVPVPSA
             ISAGNDVNKLIISGQSSKXKDEVEVNEIIEAGFNWKNDFIPKPYHDTGTVMSAIT
             QEGLRLGVVRPEDFVAELTRPEOFQIAPLVKGNFPFGRWYHRSNWAHAR
             NALVVAAREAQRLGVRFVAGSPQGRVITLIFENNDRKAVTADGKIWRABQTLICAGA
             AAGQFLDFKDLRLPTAWTLVHIQLKPEERAQYKKNMPVVFVNIKGFPEPDERGEIKI

```

FEATURES

```

source
1. .1314
/mol_type="mRNA"
/isolate="GPI"
/db_xref="taxon:33178"
/cell_line="E.coli SOLR"
1. .1314
/gene="faoA"
1. .1314
/gene="faoA"
/codon_start=1
/product="fructosyl amino acid oxidase"
/protein_id="CAA70218.1"
/db_xref="GI:1653970"
/db_xref="GOA:Q92204"
/db_xref="UniProt/TsEMBL:Q92204"
/translation="MPVTKSSSILLIGAGTWGCGSTALHLARCYTNVTLDPVPVPSA
ISAGNDVNKLIISGQSSKXKDEVEVNEIIEAGFNWKNDFIPKPYHDTGTVMSAIT
QEGLRLGVVRPEDFVAELTRPEOFQIAPLVKGNFPFGRWYHRSNWAHAR
NALVVAAREAQRLGVRFVAGSPQGRVITLIFENNDRKAVTADGKIWRABQTLICAGA
AAGQFLDFKDLRLPTAWTLVHIQLKPEERAQYKKNMPVVFVNIKGFPEPDERGEIKI

```

ORIGIN

```

Query Match      48.3%; Score 685; DB 8; Length 1314;
Best Local Similarity 70.7%; Pred. No. 5.7e-192;
Matches 926; Conservative 0; Mismatches 380; Indels 3; Gaps 1;

QY      38 CGTCACTAAGTCATCATCTCTCTGATCGTTTGGTGGCGGACCTTGGGCGACCTCAACGG 97
Db      5 CAGTCAACCAAGTCTTGTGATATTGATCATCGGGCGGGCACCTGGGGTGTCTCAACG 64
QY      98 CTCTGACCTCTCGCGCGCGCGATATACCAAGTGTACCGTGTGACACCGCTTCCTGTGTC 157
Db      65 CCTGTGATCTTGGCGCAGAGGATACCAATGTCACTGTCTTGAACCGCTACCGGTTTC 124
QY      158 CTAGCGCCATCTCCGCGGAAACGAGTGAACAAAGTCAATTAGAGTGGCGCAATATCGA 217
Db      125 CATCAGCCATTTCCGCGGCAACGAGTCAACAAGATCATCTCGTCCGCGCAGTACAGCA 184
QY      218 ATAAACAAGACCAATCGAAGTGAATGAGATCTTGGCGGAAGGCGTTTAACGGTTGA 277
Db      185 GCAAGAAGGACGAGGTGCAAGTCAATGAGATTATCGCCGAACAGGCTTCAATGGCTGA 244
QY      278 AGAACGACCCGCTTTTCAAAACGTTATTATCATGATACGGCGCTGCTGATGTCTGTGCT 337
Db      245 AAATGACCCCATCTTCAAGCCGTACTACCAAGACCGCGTCTGTGATGTCCGCCACCA 304
QY      338 CGCAGAGCGGCTTGATTCGCTGGGCTCCGGTACGTCGGGCGAGGATCTTAATCTGG 397
Db      305 CACAGGAAGGATTGGAGCGTCTGGGGTCCGGTGCAGCTGAAAGATGAACCCGATGTAG 364
QY      398 TGAACCTTACCGCGCGAGCAATTCGTAACCTGCCCCCGGAAGCGGTGTTCAGAGTG 457
Db      365 CGGAATTGATCTCGCGCGGAGCAGTTCCGCGAGCTGGCCCCC---GGCGTCTTGAAGGGTA 421
QY      458 ATTTTCGCGGTTGGAAGGGTACTTTTGGCGCTTCGCGCGCTGGCTGGGCAATGCAAGGA 517
Db      422 ACTTCCCGGTTGGAGGGGTACCAATTCGCTCAACCGGGGCTGGGCGCATGCGCGCA 481
QY      518 ATGCTTTAGTGGCAGCAGCACGGAAGCAGCGCATGGGTGTAAATTTGTTCTGGCA 577
Db      482 AGCCCTGTGTGCGCGCGCGGAGGCAAGCGCTGGGTGTGCGTCTCGTGGCGGAT 541
QY      578 CCGCGAGGGTGTGTAGTCAAGTAACTCTTGAATAAACGATGTAAAGGTGCGGTTA 637
Db      542 CGCGCAGGGCAGAGTCATCAGTTGATTTTGAAGAACCAACGATGTGAAGGGTGCGGTCA 601
QY      638 CGGGCGATGGCAAAATTTGGAGAGCGGAACGTACATTCCTGTGTGGGGCTAGCGCGG 697
Db      602 CGCGCAGCGCAAGATCTGGCGGCGGAGCAGACTATCTCTCGCTGGTGGCGCGCGCG 661
QY      698 GTCAGTTCCTAGATTTCAAGAATCAACTTCGACCAACCGCTTGGACCCCTGGGTACATTTG 757
Db      662 GCCAGTTTCTGGATTCAAGGACCAACTGCGTCCCACTGCGTGGACTCTGGTCCACATCC 721
QY      758 CGTTAAAACCGGAGAACGTTGCTGTACAAATAATACGGTTATCTTTAAATCGAAC 817
Db      722 AGTTGAAGCGGAAGAGCGTGCACAGTATAAAAACATGCGCGGTGTTCTTCAACATCGA 781
QY      818 GGGGGTTTTTTCTTTGAACCCGATGAGGAGCGCGGTGAGATTAAATATGCGATGAACACC 877
Db      782 AGGGGTTCTTCTTCGAGCGCGATGAGAGGGTGTGTAATCAAGATCTGCGACGAACACC 841
QY      878 CGGGCTACACAAATATGGTCCAGAGTGCAGACCGCAGATGATGAGATTCCTGTTGCAAA 937
Db      842 CCGGGTACAGAAATATGACCAACCGGGCGCGACCGCGCGGTGAGGAGCATTTCCCTTCGAGA 901
QY      938 AAACCCAGATTCCAAAGAGCGCAACCGCGTTCGGGCGCTTGGGCGCTGTAAGAGACATGC 997
Db      902 AGACGAGGTTCTCGAGAGGCGGAGATGCGCGTCCGCAAGCTTCTGTCTGTAACGATGC 961

```



```

Qy 998 CCAGCTGGCAGACCGTCCATTTCAGCTTCGACGCAATTCCTGGTGTCCGATACCGGA 1057
Db 962 CTCAGCTTGGAGACCGCGCTTCAGTTTCGCAAGGATCTGCTGGTGTCCGATACCGCCA 1021
Qy 1058 ATCCGGAATTCCTGATAGATCGACATCCGCAATACCAAGTCTTGTGTGGTGTGGTG 1117
Db 1022 ATCCGGAATTCATTCATGACCGTTCATCCGAATACCGTCCGTTGTTTGGTGTGGTG 1081
Qy 1118 CGAGCGAAGAGGTTTAAATATCTGCTTCTATTTGGGAATCTCATTTGTTGACCGATGG 1177
Db 1082 CTTGAGCAGCAGGCTTCAAAATATCTCCCTCGATCGGAAGCATCATCGCAGACCCATGG 1141
Qy 1178 AAGTAAAGTCCGCAAAAATTCACGAATTAATCAAGTGAACCCCGACATTCGCGCGA 1237
Db 1142 AGACAACACCCCGCAAAAATTCACAAAGTGAATCCGCTGAGCCCGCAATCCGATCA 1201
Qy 1238 ACCGTAAGTCCGCTGATCTCTGGGGGTTTGGCGGTTCAAAATCGTGTGATGATTTTC 1297
Db 1202 ACCGTAAGTCCGCGGACAGATTAGTTCGATTTGGAGGCGCCACCGGCTCATGGATTCA 1261
Qy 1298 ATGATGTAAGGAATGACCAATGTTTCAGTATCTGATATTTTCAAGCT 1346
Db 1262 ATGAAGTGAAGGAGTGAATTAATTCACCAAGGAGCATCTCGAAGTT 1310

RESULT 7
AB180733
LOCUS
DEFINITION Aspergillus oryzae FAOAO2 mRNA for fructosyl-amino acid oxidase,
complete cds.
ACCESSION AB180733
VERSION AB180733.1 GI:53850451
KEYWORDS
SOURCE Aspergillus oryzae
ORGANISM Aspergillus oryzae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE
1 Akazawa,S., Karino,T., Yoshida,N., Katsuragi,T. and Tani,Y.
Functional Analysis of Fructosyl-Amino Acid Oxidases of Aspergillus
oryzae
Appl. Environ. Microbiol. 70 (10), 5882-5890 (2004)
JOURNAL
PUBMED 15466528
REFERENCE
2 (bases 1 to 1311)
Akazawa,S., Karino,T., Yoshida,N., Katsuragi,T. and Tani,Y.
Direct Submission
Submitted (29-MAY-2004) Shin-ichi Akazawa, Graduate School of
Biological Sciences, Nara Institute of Science and Technology,
8916-5 Takayama, Ikoma, Nara 630-0192, Japan
(E-mail: s-akazawa@bs.nara.ac.jp, Tel: 81-743-72-5423,
Fax: 81-743-72-5429)
FEATURES
             source
             1..1311
             /organism="Aspergillus oryzae"
             /mol_type="mRNA"
             /db_xref="taxon:5062"
             1..1311
             /gene="FAOAO2"
             1..1311
             /codon_start=1
             /gene="FAOAO2"
             /product="fructosyl-amino acid oxidase"
             /protein_id="BAD54825.1"
             /db_xref="GI:53850452"
             translations="MTVAKSSSILIGAGTWGASTALHLGRGVNTVLDPVTPSA
             ISAGDVNKLIISQGSNKKDEIEVNEILAEAFKWTTPDKPYHDTGVMSACS
             SAGDRUGIRVEEEDVSEVTKPFHQALPAVLKGNFPFGMRGVRHNSNAGWAHAR
             NALVAIAREAKLGKVFPTQGRVITLI FENNDRKGVANTADGKI WRAEQTIVLCAGN
             AAQFLDFKQDLRTAWTLAHLRKPEERALKYLPVIFNTEKGFPEPDERGEIIC
             DEHPGTNMVMSADHLTSLPFEKTOIPKSEARVALLSETWPLADRPFSFARVCW
             CADTANREFIIDRHPESLVLGGSGRGFKYLPISGLNLDVAIEDKVEKVKHLTR
             WSPDIADVRKWRDTLGRFPGPNRUMDFHVKWNTVNQKDTAKL"
ORIGIN

```

```

Query Match 46.3%; Score 657.6; DB 8; Length 1311;
Best Local Similarity 69.9%; Pred. No. 8.1e-184;
Matches 917; Conservative 0; Mismatches 389; Indels 6; Gaps 2;

Qy 38 CCGTTCATAGTCATCATCTCTCTGATCTGTTGGTCCGGGACTTTGGGGACCTTCAACGG 97
Db 5 CTGTGCGCAAAATCTTCTCAATCTCTGATCATTCGGCGCAGGCACCTTTGGGGCGCTTCAATG 64
Qy 98 CTCTGCACCTCGCGCGCGCGGATATACCAACGTTACCGTCTGAGACCCCTATCTGTGCC 157
Db 65 CCCTTCACCTTGGTTCGAGAGGATACACCAATGTCAACGCTTAGACCCCTTACACAGTGC 124
Qy 158 CTAGCGCATCTCGCGCGGAAAACGAGTGAACAAAGTCATTAGCAGTGGCCAAATATTGGA 217
Db 125 CCTCAGCAATTTTCAGCTGGAAATGACGTGAACAAAGATCATCTCTCGGGGCAATACAGCA 184
Qy 218 ATAAACAAAGACGAAATCGAAGTGAATGAGATCTTTGGCGGAAGAGCGCTTTTAAACGTTGGA 277
Db 185 ACAAAGAGATGAGATTGAAGTTAAACGAATTTCTCGCCGAGGAGGCAATTCAGAGCTGGA 244
Qy 278 AGAACGACCCGCTTTTCAAAACCGTATTATCATGATACGGGCGCTGCTGATGTCTCTGTCT 337
Db 245 CAACCGACCTTTGTTCAAGCCATACTACCAAGACACTGGCGTTGTAATGTCTGCTTGA 304
Qy 338 CGCAGGAGGCTGATTCGCTCGGCGTCCGGGTACGTCGCGGCGAGGATCTTAATCTGG 397
Db 305 GCAGCGCGCTCTGGATTCGCTCGGAATTCGAGTAAGCGCGGAGAGGAACTGATGTTT 364
Qy 398 TGGAACTTACCCCGCGGAGCAATTTCTGTAACCTGGCCCGGAGGCGGTGTTGCAAGGTG 457
Db 365 CCGAAGTCAGAAAGCCGAGGACATCTCCGCCAATCGCCCCCGCT---GTGCTGAAGAGAA 421
Qy 458 ATTTTCCGGGTGGAAAGGGTACTTTTCCGCGCTTCGCGCGTGGCTGGGCGACATCAAGGA 517
Db 422 ACTTCCGGGTGGAGAGGCTACCAATCTGTTTCGAACGCTGGCTGGGCGCACGCCGAA 481
Qy 518 ATGCTTTAGTGGCAGCAGCAGCGAAGCAAGCGCATGGGTGTAAATTTGTTACTGGCA 577
Db 482 ATGCCCTCTGGTGGCGCTTATACGCGAAGCAGAGAAACTTGGTGTAAATTCGTAAACGCA 541
Qy 578 CCCCGAGGCTCTGTAGTCACGTTATCTTTCGAATAACGATGTAAGAGTGCCTGTTA 637
Db 542 CC---CAAGGAAGAGTCATCACCTTATCTTCGAGAACACAGCTCAAGGGCGCAGTCA 598
Qy 638 CGGCGCATGCGCAAAATTTTGAGAGCGGAACTGATCTCTGTGTGGTGGGCTAGCGCGG 697
Db 599 CCGCCGACGGAAGATCTGCGCGCGGAGCAACAGTTCTCTGCTGGCTGGCCTGCTG 658
Qy 698 GTCAGTTCTTAGATTTCAAGAAATCAACTTCGACCAACCGCTTGGACCCCTGGTACACATTG 757
Db 659 CGCAGTTCTTGGATTTTAAAGGACAGCTCGCGCCCAACGCGATGGACACTCGCCCATATCC 718
Qy 758 CGTTAAACCGGAAGAACGCTGCTGTGTACAAAATATACCGGTTATCTTTAAACATCGAAC 817
Db 719 GGCTCAAACTGAGGAACCGCGCTCTACAAAACTTTCGCCCGTGAATTTCAACATTTAGA 778
Qy 818 GGGGGTTTTTCTTTGAAACCGATGAGGAGCGCGGTGAGATTAAATAATTCGATGAACAC 877
Db 779 AAGGATTTTCTTCGAGCTGTATGAGGAGCGCGGGGAGATCAAGATCTGCGACGAACATC 838
Qy 878 CGGGCTACACAAATATGGTCCAGATGTCAGACGCGCAGATGATGAGCATTCCTGTTGAAA 937
Db 839 CGGGATACACTAACATGGTTAAATCTCGCGATGGCCACTTGAACGAGTTTGGCCCTTTGAGA 898
Qy 938 AAACCCGATTTCCAAAAGAGCGGAACCGCGCTTCGGGCCCTCTCTGAAAGAGCAATTC 997
Db 899 AGACCCAGATTTCCCAAGGAGCTGGAAGCTAGAGTTCAGAGCTTTACTATCGGAGACCATGC 958
Qy 998 CCCAGCTGGCAGACCGTCCATTTCAGCTTTCGACGCAATTTGCTGTGTGGCGGATACCGCGA 1057
Db 959 CTCANTTAGCCGATCGCCCATTTAGCTTTCGCCCGGCTTCTGCTGGTGTGGGACACCGCA 1018

```



```
Db 1213 GAGTAAACTACAGAAGGAGGTGAAGGACATCGTTTCGATGGAGGCCAGAGACGGCTGTGCA 1272
Qy 1239 CCGTAACTGGCGTGATCTCTGGGGCGTCTTTGGCGGTCCTCAAAATCGTGTGATGATTTTCA 1298
Db 1273 TCGAGATTGGAGCGCACTCAGATCGCTTTTGGCGGCGCTGACAGGATCATGGATTTTCA 1332
Qy 1299 TGATGT 1304
Db 1333 GCAGGT 1338

RESULT 9
LOCUS AB180732 1338 bp mRNA linear PLN 07-OCT-2004
DEFINITION Aspergillus oryzae FAOa01 mRNA for fructosyl-amino acid oxidase,
complete cds.
ACCESSION AB180732
VERSION AB180732.1 GI:53850449
KEYWORDS
SOURCE Aspergillus oryzae
ORGANISM Aspergillus oryzae
Eukaryote; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1
AUTHORS Akazawa, S., Karino, T., Yoshida, N., Katsuragi, T. and Tani, Y.
TITLE Functional Analysis of Fructosyl-Amino Acid Oxidases of Aspergillus
oryzae
JOURNAL Appl. Environ. Microbiol. 70 (10), 5882-5890 (2004)
PUBMED 15466528
REFERENCE 2 (bases 1 to 1338)
AUTHORS Akazawa, S., Karino, T., Yoshida, N., Katsuragi, T. and Tani, Y.
TITLE Direct Submission
JOURNAL Submitted (29-May-2004) Shin-ichi Akazawa, Graduate School of
Biological Sciences, Nara Institute of Science and Technology,
8916-5 Takayama, Ikoma, Nara 630-0192, Japan
(E-mail:s-akazawa@bs.ais.t-nara.ac.jp, Tel:81-743-72-5423,
Fax:81-743-72-5429)
FEATURES
source
1. .1338
/organism="Aspergillus oryzae"
/mol_type="mRNA"
/db_xref="taxon:5062"
1. .1338
/gene="FAOa01"
/codon_start=1
/product="fructosyl-amino acid oxidase"
/protein_id="BAD54824.1"
/db_xref="GI:53850450"
/translations="MTSSKLTPTSSILIVGATGWCSTALHLARRGKYNVTVDLPHPV
PSPIAAGNDINKIMEHREVKASSETDPMIAFSTCTRAALKGWNKDPVFPYFHTGAI
VSGHTASLIKHIOBHEIDSDAGFIIKLNTAEDFRKTPPGILTGNFPKWLKNGA
GWIHAKMESAETEARLGVFTITGSPGDVSLIYENGVDVVGATADGTVRAOHT
ILSAGSDRLHFFKQLRPTANTLCHIEMTPBEAKKYNLPLFNVAAGFFPMEPDED
NHELKICDEHPGVNFPDPKHGEVNSIPFAKHQIPLBAEARDFLRDTPHPLADR
PLSPFARICDADVTDFRAFLDRHPEYRSLSLAVGGSGNAGQMPTTGGTADALEGNI
QKELKHALRWRPBTAAQRDKWDTQNRFGGPNKVMDFQVGENEWTIKIGKSRL"
ORIGIN
Query Match 24.78; Score 350.4; DB 8; Length 1338;
Best Local Similarity 55.94; Pred. No. 2.1e-92;
Matches 713; Conservative 0; Mismatches 551; Indels 12; Gaps 2;
Qy 41 TCACTAAGTCATCATCTCTCTGATCGTTGTCGCGGACCTTGGGGCACCTCAACGGCTC 100
Db 17 TGACTCCACATCATCTATCTTAATGTTCGTGCGAGGACCTGGGGTGTCTACTGCTT 76
Qy 101 TGCACCTCGCGCGCGGATATACCAACGTTACCGTGTGGACCCCTATCTGTCCCTA 160
Db 77 TACATCTTGGCCGTCGAGGATACAAAATGTACGGTCTTAGATCCGACCCGCTCCTT 136
Qy 161 GCGCCATCTCGCGGGGAAACGAGCTGAACAAAGTCAT-----TAGCAGTGGCCAATATT 214
```

```
Db 137 CTCCCATGACGTGGCAATGACATTAACAAGATTATGGAGCAAGGAGTAAGAAGCTT 196
Qy 215 CGAATAACAAAGACGAAATCGAAGTGAATGAGATCTTTGGCGGAAGAGCGCTTTAACCGTTT 274
Db 197 CTGAAACCGATCTCTTGGAGTATCGCTTCTCAAACATGACACAGAGCTGCACTGAAAGGTTT 256
Qy 275 GGAAGAAGCAAGCCGCTTTTCAAACCGTATTATCATGATACGGGCTGCTGATGCTGTCTT 334
Db 257 GGAATAACGACCCAGTATCTCCAGCCATATCTCCATGAACCGGGGCAATAGTTTCTTGCC 316
Qy 335 GCTCGCAGGAGGGCTCGATCGCTGGCGTCCGGGTACGTCGCGGCGAGGATCCTTAATC 394
Db 317 ACACCGCTCTTTGATTAAACATATACAGAACAGAAATCGACTCTGTCAGACCGCGAGT 376
Qy 395 TGGTGGAACTTACCCGCCCGGAGCAATTTTCGTAAACTGGCCCGGGAAGGCGTGTGTCAAG 454
Db 377 TCATAAAATTTGAACACCGCAGAGGATTTCCGCAAAACTATGCCCCCGGGAATCCTCACTG 436
Qy 455 GTGATTTTCCGGGTTGGAAGGGTACTTTTGGCGGCTCCGGCGCTGGCTGGGCACATGCAA 514
Db 437 GCAACTTTCCCGCTGGAAGGGCTGGCTGAACAAGACCGGCGCGGATGGATCCACGCCA 496
Qy 515 GGAATGCTTAGTGGCAGCAGCAGCGAAGCAGCGCATGCGGTGTAATAATTTGTTACTG 574
Db 497 AGAAGGCCATGTTCTCCGCATACACCGAAGCAAGCGCTTAGGAGTCACTTTCAATCACCG 556
Qy 575 GCACCCCGCAGGGTCTGTAGTCACGTTAAATCTTTGAAAAATAACGATGTAAAAAGTGC 634
Db 557 GCTCCCTCGAAGGAGAGGTTGTATCTCTAATTTACGAGAATGGAGACGTAGTCGAGCCA 616
Qy 635 TTACGGCGGATGGCAAAATTTGGAGAGCGGAACGTACATCTCTGTGTCTGGGGCTAGCG 694
Db 617 GAACGGCGCAGCGGACCGTCCACCGAGCAGACCATACCATTTCTTCCGCGAGGGCTGSCA 676
Qy 695 CGGTCAGTTTCTTAGATTTTCAAGAAATCAACTTGGACCAACCGCTTGGACCTGTGTACACA 754
Db 677 GTGATGCTCTCTGGACTTTAAGAAAACGCTCCGCTCTCCGCTGGACGCTCTGCCACA 736
Qy 755 TTGCGTTTAAAAACCGGAAGAACGTGCGTTGTACAAAAATATACCGGTTATCTTTAACATCG 814
Db 737 TCAGAAATGACGCCCGAGCGGCCAAGAGTACCGGAATCTTCTGTGCTGTTCAACGTCG 796
Qy 815 AACGGGGGTTTCTTTGAAACCCGATGAGAGCGCGGTGAGATTTAAATAATATGCGATGAAC 874
Db 797 CTAAGGGGTTCTTTCATGGAACCTGATGAGGATAATCATGAGCTTAAGATCTGCGACGAGC 856
Qy 875 ACCGGGCTACACAAATAT-----GGTCCAGAGTGCAGACGCGACCGATGATGAGCATTC 928
Db 857 ATCTTGATATTGCAACTTCTGTCGCGACCCGGAAGCACGCGGCTGAGGTGGCAGTATCC 916
Qy 929 CGTTTCAAAAAAACCCAGATTCCAAAAAGAACCGCAAAACCGCGCTTCCGGGCCCTGCTGAAAG 988
Db 917 CATTTGCAAGCATCAGATCTCTTGAAGCGGAGGCCCGTGGCAAGGACCTTCTCCGTG 976
Qy 989 AGAACAATGCCCCAGCTGGCAGACCGTCCATTCAGCTTTCGACGCGATTTGCTGGTGTGCGG 1048
Db 977 ATACGATGCTCATCTTGTGTCGACCACTGTCTTTTGTCTGCTATATGCTGGGATGCTG 1036
Qy 1049 ATACCGGCAATCGGAATCTCTGATAGATCGACATCCGACCGATACCAAGTCTTCTGTGTTG 1108
Db 1037 ATACAGTGGATCGCGCTTCTTGATGATAGGCAATCTCGAGTATCGCTCTTTACTGCTTG 1096
Qy 1109 GCTGTGTGTCGAGCGGAAGAGGGTTTAAATATCTGCGCTTCTTATTTGGGAATCTCATTTGTTG 1168
Db 1097 CTGTCGCTGGATCTGTTAATGAGAGCCATGCAATGCTTACCATTTGGTGGTTCATAGCG 1156
Qy 1169 ACAGCATGGAAGTAAAGTCCCGCAAAATATTCAGAAATTAATCAAGTGAACCCGGACA 1228
Db 1157 ATGCTCTGGAGGGAACCTGCAAAAGGAATCTGAAGCATGCACTACGCTGGAGGCTCGAGA 1216
Qy 1229 TTGGGCGAACCGTAACTGGGCTGATACTCTGCGGGCGTTTGGCGGTCCTCAATCTGTGA 1288
```

Db 1217 TTGCCGCCCAACGAGACTGGAAGGATACGCAAAATAGATTCCGAGGTCCGAATAAAGTAA 1276

QY 1289 TGGATTTTCATGATGT 1304

Db 1277 TGGATTTCCAAAGGT 1292

RESULT 10

LOCUS BD173827 1320 bp DNA linear PAT 18-FEB-2003

DEFINITION Compositions for assaying glycoprotein.

ACCESSION BD173827

VERSION BD173827.1 GI:28415160

KEYWORDS WO 02061119-A/1.

SOURCE unidentified

ORGANISM unidentified

REFERENCE 1 (bases 1 to 1320)

AUTHORS Kozuma,T., Yoshioka,I., Arai,M., Sumitani,J. and Imamu,S.

TITLE Compositions for assaying glycoprotein

JOURNAL Patent: WO 02061119-A 1 08-AUG-2002;

ASAHI KASEI CORP, TAKUJI KOZUMA, ISSEI YOSHIOKA, MOTOO ARAI, JUNICHI SUMITANI, SHIGEYUKI INAMURA

OS Fusarium oxysporum IF0-9722

PN WO 02061119-A/1

PD 08-AUG-2002

PF 30-JAN-2002 WO 2002JP000721

PR 31-JAN-2001 JP 01P 022953, 16-FEB-2001 JP 01P 039796 PR

08-AUG-2001 JP 01P 240002

PI TAKUJI KOZUMA, ISSEI YOSHIOKA, MOTOO ARAI, JUNICHI SUMITANI, PI SHIGEYUKI INAMURA

PC C12Q1/26, C12Q1/37, C12N15/09, G01N33/68, G01N33/72 CC

Compositions for assaying glycoprotein

EH Key Location/Qualifiers

FI CDS (i). (1320).

FEATURES

source 1..1320

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

ORIGIN

Query Match 23.6%; Score 335.2; DB 6; Length 1320;

Best Local Similarity 55.3%; Pred. No. 7.2e-88;

Matches 721; Conservative 0; Mismatches 568; Indels 15; Gaps 3;

QY 41 TCACTAAGTCATCATCTCTCTGATGCTTGGTGGCGGACTTGGGGCACTCAACGGCTC 100

Db 11 TCACCAACAGCTCCCAAAATCTCATCTGGTGGCGGAACCTTGGGGATGCTCAACTGCC 70

QY 101 TGCACCTCGCGCGCGGATATACCAACGTTACCGTGTGGACCCCTATCTGTCCCTA 160

Db 71 TCCATCTCGCCGTGGGGTTACCAACGTCATCTGTTCTCGATGTCAATCGCATCCGT 130

QY 161 GCGCCATCTCGCGCGGAACGAGCTGAAACAAAGTCATTAGCAGTGGCCAAATATTCGAATA 220

Db 131 CACCGATATCAGCGCGGCATGATGTAAACAACTTGCTGCCGACATGCTCGACTGCCGATA 190

QY 221 ACAA-----GAGAAATCGAAGTGAATGATCTTGGCGGAAGAGCGGTTAAACGGTT 274

Db 191 GCAAAGGTGATGATGAAGACTCAATCTGGAAGCACTTAGCTACGCCGCACTCAAGGAT 250

QY 275 GGAAGACGACCGCTTTTCAACCGTATTATCATGATACGGGCTGCTGATGCTGCTT 334

Db 251 GGCTCACGACCCCTGCTTCCAAACCATCTGCCAATAACAGGCTGTGCTGGTGGCT 310

QY 335 GCTCGCAGGAGGGCTGGATCGCTGGCGCTCGCGGTACGCTCGGGCGAGGATCCTAATC 394

Db 311 CAACACCAAGTCTATCAACGAGCTGTAGAGATGAGATCGGTGACGACATCGACCACT 370

QY 395 TGGTGGAACTTACCGCCCGGAGCAATTCGTAAATCGGCCCGGAAAGCGGTGTGCAAG 454

Db 371 ATACACCTCTCAACACAGCAGAAGATTTTCAAGAAAGACCATGCGCTGAGGGTATCCTGACAG 430

QY 455 GTGATTTTCCGGGTTTGGAAAGGTAATTTCGCGGTTTCGCGGCTGGCTGGGCACATGCAA 514

Db 431 GTAACTTTCCAGGCTGGAAGGGCTTTTCAAGCCCAAGGGTCTGCTGGTTGATGCTC 490

QY 515 GGAATGCCCTTAGTGGCAGCAGCAGCGAAGCAGCGATGGGTGTAATAATTTGTTACTG 574

Db 491 GAAAAGCTATGAAAGCTGCTTTTGAAGAGAGCAGAGGCTTGGTGTCAAAATTCATCACTG 550

QY 575 GCACCCCGCAGGTCGTGTAGTCACCTTAATCTTTGAAAATAACGATGTAAGAGGTGCCG 634

Db 551 GCTCTCCGAAAGAAAGGTGGAGAGTCTGATCTTTGAAGACGGCGATGTTTCGAGGTGCCA 610

QY 635 TTACGGCGCATGGCAAAATTTGGAGAGCGGAACGTCACTTCTCTGTGTGTGGGGCTAGCG 694

Db 611 AGACGGCAGATGTTAAGGAGCAGACAGCGGATCGAATCTTCTTCGCTGGTGTCTCAG 670

QY 695 CGGGTCAGTTCTTAGATTTCAAGAATCAACTTCGACAAACCGCTTGGACCTGGTACACA 754

Db 671 CAGAGTTCTTCTCGATTTTGAAGAACAGATCCAGCTACGGCGTGGACCTGGGCCATA 730

QY 755 TTGCGTTAAACCGGGAAGACGTCGCTGTGTACAAAATATACCGGTTATCTTTAAACATCG 814

Db 731 TCAGATGACACCCAGAAAGAACCAAGCTGTACAAAGAACTTGCACCTCTTTTCAACATCA 790

QY 815 AACGGGGTCTTTTCTTTGAACCCGATGAGGAGCGCGTGAAGTTAAATAATATGCGATGAAC 874

Db 791 ACCAAGTTCTTCATGGAACCTGTATGAGGATCTTTCATCAACTCAAGATGTGCGATGAAC 850

QY 875 ACCGGGCTACACAATAATATGTTCCAGATGTCAGACGGCAGC---ATGATGAGCATTCCTG 931

Db 851 ATCCGGGCTACTGCAACTGGGTTGAAAAGCTGCTTTCTAAGTACCCCCAGTCCATCCCT 910

QY 932 TCGAAAAACCCAGATTCCAAAGAGCGGAAACGCGCTTCGGGCCCTGCTGAAAGAGA 991

Db 911 TCGAAAGCATCAAGTGCCAAACCGAGCTGAACGAGCATGAAGCAGTTTCTGAAAGATA 970

QY 992 CAATGCCCCAGCTGGCAGACCGTCCATTCAGCTTTCGACGCAATTTGCTGGTGTGCCGATA 1051

Db 971 TCATGCTCAGCTTGCAGATCGGCCCTGTTTCATCTCGAATCTGCTGGTGGCTGATA 1030

QY 1052 CGCGAATCGGAATTCCTGTATAGATCGACATCCGAGTACCAACAGCTTGTGTGGGCT 1111

Db 1031 CACAGATAGAATGTTCTCTGATCACCTATCATCTCGACATCCCTCATTGTCTGCTT 1090

QY 1112 GTGGTCGAGCGGAAGAGGGTTTAAATATCTGCTTTCTATTGGGAATCTCATTTGTCACG 1171

Db 1091 CAGGTGATTCGGCAGCGGTTACAGCATATCATCAATTTGGAAGTTTCTCTGACT 1150

QY 1172 CGATGGAAGTAAAGTTCGCCCAAAAAATTCAGAAATTAATCAAGTGGAAACCCGAGCATTTG 1231

Db 1151 GTATGAGGTTACGCTTGAGGAAGGTTTGCCTAAGTTCTGGAGATGGCGACAGAGAAT 1210

QY 1232 CGCGAAGCGTAACTCGGCTGATGATCTCGGGGCTTTTGGCGGTTCGAATCTGTGATGG 1291

Db 1211 TTACCGAGTTCTGGGTAAAGATCTCTCGATCGGTTGGAGCTGACGATGAAGATCATGG 1270

QY 1292 ATTTTC-----ATGATGTGAAGAAATGGACCAATGTTCAAGTAT 1329

Db 1271 ATTTGCCAAGAGTGTATGAGGGATGGCAAAATATCAAGAAAT 1314

RESULT 11

LOCUS E16562

DEFINITION Fusarium oxysporum gene for fructosylamine oxidase.

ACCESSION E16562

VERSION E16562.1 GI:5711245

KEYWORDS JP 1998201473-A/1.

SOURCE unidentified

ORGANISM unclassified.

REFERENCE 1 (bases 1 to 1320)

AUTHORS Shimizu,A. and Koga,S.
TITLE FRUCTOSYL AMINE OXIDASE-PRODUCTIVE SUBSTANTIALLY PURE ORGANISM
JOURNAL ASAHII CHEM IND CO LTD
COMMENT OS Fusarium oxysporum
PN JP 1998201473-A/1
PD 04-AUG-1998
PF 20-JAN-1997 JP 1997007101
PI SHIMIZU AKIRA, KOGA SHINJI
PC C12N15/09,C12N1/21,C12N9/06//C12Q1/26,(C12N1/21,C12R1/19), PC
(C12N9/06,
PC C12R1/19);
CC strandedness: Double;
CC copology: linear;
FH Key Location/Qualifiers
FT source 1..1320
FT /strain='IFO-9972';
FT mat_peptide 1..1320
FT /product='fructosylamine oxidase'.
FEATURES
source Location/Qualifiers
1..1320
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'
ORIGIN
Query Match 23.68; Score 335.2; DB 6; Length 1320;
Best Local Similarity 55.38; Pred. No. 7.2e-88;
Matches 721; Conservative 0; Mismatches 568; Indels 15; Gaps 3;
41 TCACTAAGTCATCATCTCTCTGATCGTGTGGCGGACCTTGGGGACCTCAACGGCTC 100
11 TCACCAACAGTCCCAATTTCTCATCTGTGGTGGGAACTTGGGGATGCTCAACTGCC 70
101 TGCACCTCGCGCGCGGATATACCAACGTTACCGTGTGGACCCCTATCTGTCCCTA 160
71 TCCATCTCGCGCGCGGTTACACCAACGTCATCTTCTCGATGTCAATCGCATCCCGT 130
161 GCGCCATCTCGCGGGAACGAGTGAACAAAGTCATTAGCNGTGGCCAAATATTCGATA 220
131 CACCGATATCAGCGGGCATATGATGAACAACTTGTGCGGCACTGTGCACTGCGGATA 190
221 ACAA-----GACGAATCGAAGTGAATGAGATCTTGGCGGAAGAGCGGTTAAACGGTT 274
191 GCAAGGTGATGAAGACTCATCTGGAAGCACTTAGCTACCGCGAGCTCAAGAT 250
275 GGAAGAACGACCGCTTTTCAACCGTATTATCATGATACGGGCTCTGATGTCTGCTT 334
251 GGCTCCACGACCTGTCTTCAACCACTTCTGCCACAATACAGGCTCTGTGCTGGCTGGCT 310
335 GCTCGAGGAGGGCTGTGATCGCTGGCGTTCGGGTACGTCGGGGGAGGATCCTAATC 394
311 CAACACCAAGTCTATCAAGCAGCTGTGAGAAGATGAGATCGGTGACGATCGACCACT 370
395 TGGTGAACCTTACCGCGCGGACCAATTCGTAACTGGCCCGGAGGCGGTGTCGAAG 454
371 ATACACCTCTCAACAGCAGAGATTTGAGAAGACCATGCTGAGGGTATCTGACAG 430
455 GTGATTTTCGGGTGGAAGGGTACTTTGCGGCTTCCGGCTGGCTGGGCAATGCAAA 514
431 GTAATTTTCAGGCTGGAAGGGCTTTTCAAGCCACGCGGTCTGTGTTGGGTTTCATGCTC 490
515 GGAATGCTTAGTGGCAGCAGCAGCGGATGATGATGATGATGATGATGATGATGATGATG 574
491 GAAAGAGCTATGAAGCTGCTTTTGAAGAGAGCGAGAGGCTTGGTGTCAAAATTCATCACTG 550
575 GCACCGCGAGGGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 634
551 GCTCTCCGAGGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 610
635 TTAGGGCGATGGCAAAATTTGGAGAGCGGAACGATACATCTCTGTGTGCTGGGGCTAGCG 694

Db 611 AGACGGCAGATGTAAGGACGACGACGATCGAATATTTCTTTCGGCTGGCTTCAG 670
Qy 695 CGGTCAGTTCCTAGATTTCAGAAATCAACTTCGACCAACCGCTTGGACCTCGGTACACA 754
Db 671 CAGAGTTCCTCGATTTTGAGAACCCAGATCCAGCTCGGCGTGGACCTCGGGCCATA 730
Qy 755 TTGCGTTAAACCGGAAGAACGTCGTTGTACAAAATATACCCGTTATCTTTTAACTCG 814
Db 731 TCCAGATGACACCAAGAAACCAAGCTGTACAAAGACCTGCCACCTCTTTTCAACATCA 790
Qy 815 AACGGGGTTTCTTTGAAACCGGATGAGGAGCGGCTGAGATTAAATATGCGATGAAC 874
Db 791 ACCAAGGTTTCTTCATGGAACCTGATGAGGATCTTCACTCAAGATGTCGATGAAC 850
Qy 875 ACCCGGCTACACAAATATGTTCCAGAGTGCAGACGGCAGC---ATGATGAGCATTTCCGT 931
Db 851 ATCCGGCTACTGCACTGGTTCAAAAGCTGTTTCAAGTACCCCGGCTCCATCCCT 910
Qy 932 TCAGAAACACCCAGATTCCAAAAGAACCGGCTTGGGGCCCTGCTGAAAGAGA 991
Db 911 TCGCAAGCATCAAGTGCACACCGAGGCTGAACGACGATGAAGCAGTTTCTGAAAGATA 970
Qy 992 CAATGCCCGAGCTGGCAGACCGTCCATTGAGCTTCGACGATTTGCTGTGTCGCCGATA 1051
Db 971 TCATGCTCAGCTTGCAGATCGGCCGCTTGTTCATGCTCGAATCTGCTGGTGGCTGATA 1030
Qy 1052 CCGCAATCGCAATCTCTGATAGATCGACATCGCAGTACCAAGCTCTGTGTGGCT 1111
Db 1031 CACAGGATAGATTTCTCTGATCACCCTATCATCTCGACATCCCTGATCTGCTT 1090
Qy 1112 GTGTCGAGCGGAAGAGGTTTAAATATCTGCTCTTATTTGGAATCTCATTTGACG 1171
Db 1091 CAGGTGATTCGGCAGCGGTTACAAGCATATCATCAATTTGGAAGTTCACTCTGACT 1150
Qy 1172 CGATGAGTAAAGTCCCGCAAAATTCAGATTAATCAAGTGAACCGGACATTTG 1231
Db 1151 GTATGAGGGTACGCTTTGAGGAAGGTTTGCAGGTTCTGGAGATGGCGACAGAGAT 1210
Qy 1232 CGCGAACCGTAACTGCGGTGATACCTGCGGGCTTTTGGCGTCCAAATCTGTGTGG 1291
Db 1211 TTACCGATTCGGGTAAAGATCTCTGATCGTTTGGAGCTGACGATGAATCATGG 1270
Qy 1292 ATTTTC-----ATGATGTGAAGAAATGACCAATGTTCAGTAT 1329
Db 1271 ATTTGCCAAGATGATGTAGAGGATGGAGCAAAATATCAAGAT 1314

RESULT 12

BD173829
LOCUS 1320 bp DNA linear PAT 18-FEB-2003
DEFINITION Compositions for assaying glycoprotein.
ACCESSION BD173829
VERSION BD173829.1 GI:28415162
KEYWORDS WO 02061119-A/3.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1320)
AUTHORS Kozuma,T., Yoshioka,I., Arai,M., Sumitani,J. and Imamu,S.
TITLE Compositions for assaying glycoprotein
JOURNAL Patent: WO 02061119-A 3 08-AUG-2002;
ASAHI KASEI CORP,TAKUJI KOZUMA, ISSEI YOSHIOKA, MOTOO ARAI, JUNICHI SUMITANI, SHIGEYUKI INAMURA
COMMENT OS Artificial Sequence
PN WO 02061119-A/3
PD 08-AUG-2002
PF 30-JAN-2002 WO 2002JP000721
PR 31-JAN-2001 JP 01P 022953,16-FEB-2001 JP 01P 039796 PR
08-AUG-2001 JP 01P 240002
PI TAKUJI KOZUMA, ISSEI YOSHIOKA, MOTOO ARAI, JUNICHI SUMITANI, PI SHIGEYUKI INAMURA
PC C12Q1/26,C12Q1/37,C12N15/09,G01N33/68,G01N33/72 CC

Compositions for assaying glycoprotein	
PH Key	Location/Qualifiers
FT CDS	(1)..(1320)
FT mutation	(1115).
FT	Location/Qualifiers
source	1..1320
FEATURES	/organism="synthetic construct"
	/mol_type="genomic DNA"
	/db_xref="taxon:32630"
ORIGIN	
Query Match	23.4%; Score 333.6; DB 6; Length 1320;
Best Local Similarity	55.1%; Pred. No. 2.1e-87;
Matches	720; Conservative 0; Mismatches 569; Indels 15; Gaps 3;
QY	41 TCATAGTCAATCTCTCTCATCGTTGGTCCGGACTTGGGCACTCAACGGCTC 100
DB	11 TCACCAACAGTCCCAATCTCATGTTGGTGGCGAACTTGGGATGCTCACTGCC 70
QY	101 TGCACCTCGCGCCGCGGATATACCAACGTTACCGTGTACCGTGTGGACCCCTATCTGTCCCTA 160
DB	71 TCCATCTCGCCGCTCGGGTTACCAACGTCACCTGTTCTCGATGTCAATCGCATCCGCT 130
QY	161 GCGCATCTCCGCGGAACGAGCTGAACAAAGTCAATAGCAGTGGCCAAATATTCGAATA 220
DB	131 CACCGATATCAGCCGGCATGATGTAACAAACTTGTGCGGACTGTGCGACTGCCGATA 190
QY	221 ACAA-----GACGAAATCGAAGTGAATGAGATCTTTGGCGAAGAGGGTTTAAACGGTT 274
DB	191 GCNAGGTGATGATGAAGTCAATCTGGAAGCACTTAGCTACGCCGACGCTCAAGGAT 250
QY	275 GGAAGAACGACCGCTTTTCAACCGTATTATCATGATACGGGCTCTGTGATGCTGCTT 334
DB	251 GGCTCCAGACCTCTCTTCCAAACCATCTGCCACAATACAGGCTCTGTGCTGGCTGGCT 310
QY	335 GCTCGAGAGGGCTGGATCGGCTGGGCTCGGGTACGTCGGGCGGAGGATCCTAATC 394
DB	311 CAACACCAAGTCTATCAAGCAGCTGTGAAGATGAGATCGGTGACGACATCGACGAGT 370
QY	395 TGGTGAACCTTACCGCCCGGACAAATTCGTAACCTGGCCCGGAAGGGCTGTTCGAAG 454
DB	371 ATACACTCTCAACAGCAGAAAGATTGAAAGACCATGCTGAGGTATCTCTACAG 430
QY	455 GTGATTTTCGGGTGGAAGGGTACTTTGCGGTTTCGGGCTGTGGCGGACATGCA 514
DB	431 GTAACCTTCCAGGCTGGAAGGGCTTTTACAAGCCACGGGTTCTGGTTGGGTTTCACTC 490
QY	515 GGAATGCTTAGTGGCAGCAGCAGCGGACGACGATGCGGTGTAAATTTGTTACTG 574
DB	491 GAAAGCTATGAAAGCTGCTTTTCAAGAGAGCAGAGAGGCTTGGTGTCAAAATTCATCACTG 550
QY	575 GCACCCGCGAGGCTCGTGTAGTCACGTTATCTTTGAAATAACGATGTAAAGGTGCGG 634
DB	551 GCTCTCCGGAAGAAAGGTGGAGTGTGATCTTTGAAGACGCGGATGTTGAGGGTGCCA 610
QY	635 TTACGGCGCATGTCAAAATTTGAGAGCGGAACGATACATCTCTGTGTGGGGCTAGCG 694
DB	611 AGAGGCAAGTGTAGGAGCAGAGCGGATCGAATCTTCTTCGGCTGGTCTTCAG 670
QY	695 CGGGTCAGTTCTTAGATTTCAGAAATCAATTCGACCAACCGTGTGACCCCTGGTACACA 754
DB	671 CAGAGTTCTCTCGATTTTGAGAACCCAGATCCAGCTACGGGCTGGACCCCTGGGCGATA 730
QY	755 TTGCGTTTAAACCGGAGACGTCGTTGTGACAAATATACCGGTTATCTTTAAACATCG 814
DB	731 TCAGATGACACCAAGAAACCAAGCTGTACAGAACTTGCACCTCTTTTCAACATCA 790
QY	815 AACGGGGTTTTCTTTTGAAACCGATGAGGCGCGTGAAGATTAATAATATGCGATCAAC 874
DB	791 ACCAAGTTTCTTCATGGAACCTGATGAGGATCTTTCATCACTCAAGATGTGCGATGAAC 850
QY	875 ACCCGGGCTACAAATATGTTCCAGTGCAGAGTGCAGACGCG-----ATGATGAGCATTCGCT 931
1051 ATCCGGGCTACTGCAACTGGGTTGAAAAGCCCTGGTGTCTAAGTACCCCGAGTCCATCCCT 910	
932 TCGAAAAAACCCAGATTCCAAAAGAGCGAAACGCGCTTTCGGGCGCTGCTGAAAGAGA 991	
911 TCGCAAAGCATCAAGTGCCCAACCGAGGCTGAACGACGATGAAGCAGTTTCTGAAAGATA 970	
992 CAAATGCCAGCTGGGAGACCGTCCATTGAGCTTCCAGCATTCGACGATTTGCTGGTGGCGGATA 1051	
971 TCATGCTCAGCTTGCAGATCGGCGCTTGTTCATGCTCGAATCTGCTGGTGGCGTATA 1030	
1052 CCGGAATCGCAATTTCTCTAGATCGACATCCGAGTACCACAGTCTTGTGCTGGGCT 1111	
1031 CACAGATGAAATGTTCTTGATCACCTATCATCTCGCATCCCTCACCTGCTTGTGCTT 1090	
1112 GTGGTCGAGCGGAGAGGGTTTAAATATCTGCTTTCTATTGGGAATCTCATTTGTTGACG 1171	
1091 CAGGTGATTCGGCAGCGGTTACATGATATCACATCAATTTGAAAGTTTCATCTGACT 1150	
1172 CGATGAAAGTAAAGTCCGCAAAAATTCACGAATTAATCAAGTGGAAACCCGACATTCG 1231	
1151 GTATGAGGGTACGCTTGAGGAAAGGTTTCCCAAGTTCTGGAGATGGCGACGAGAAAT 1210	
1232 CGCGAACCGTAACGCGCTGATCTCTGGGCGTTTTCGGCGTCCAAATCGTGTGATGG 1291	
1211 TTACCGAGTTCTGGGTAAGATCTCTCGATCGGTTTGGAGCTGAGCATAGATCATGG 1270	
1292 ATTTTC-----ATGATGTGAAGGATGACCAATGTTCAAGTAT 1329	
1271 ATTTGCCAAGAGTGTAGGGGATGACAAATATCAAGAT 1314	
BD173828	1320 bp DNA linear PAT 18-FEB-2003
LOCUS	
DEFINITION	Compositions for assaying glycoprotein.
ACCESSION	BD173828
VERSION	BD173828.1 GI:28415161
KEYWORDS	WO 02061119-A/2.
SOURCE	synthetic construct
ORGANISM	synthetic construct
REFERENCE	1 (bases 1 to 1320)
AUTHORS	Kozuma, T., Yoshioka, I., Arai, M., Sumitani, J. and Imamu, S.
TITLE	Compositions for assaying glycoprotein
JOURNAL	Patent: WO 02061119-A 2 08-AUG-2002;
	ASAHI KASEI CORP.TAKUJI KOZUMA, ISSEI YOSHIOKA, MOTOO ARAI, JUNICHI SUMITANI, SHIGEKUJI IMAMURA
COMMENT	OS Artificial Sequence
	PN WO 02061119-A/2
	PD 08-AUG-2002
	PF 30-JAN-2002 WO 2002JP000721
	PR 31-JAN-2001 JP 01P 022953,16-FEB-2001 JP 01P 039796 PR
	08-AUG-2001 JP 01P 240002
	PI TAKUJI KOZUMA, ISSEI YOSHIOKA, MOTOO ARAI, JUNICHI SUMITANI, PI SHIGEKUJI IMAMURA
	PC C12Q1/26.C12Q1/37,C12N15/09,G01N33/68,G01N33/72 CC
	Compositions for assaying glycoprotein
PH Key	Location/Qualifiers
FT CDS	(1)..(1320)
FT mutation	(1114)..(1115).
FT	Location/Qualifiers
source	1..1320
FEATURES	/organism="synthetic construct"
	/mol_type="genomic DNA"
	/db_xref="taxon:32630"
ORIGIN	
Query Match	23.4%; Score 332; DB 6; Length 1320;
Best Local Similarity	55.1%; Pred. No. 6.4e-87;
Matches	719; Conservative 0; Mismatches 570; Indels 15; Gaps 3;
QY	41 TCATAGTCAATCTCTCTCATCGTTGGTCCGGACTTGGGCACTTCAACGGCTC 100

Db 11 TCACCAACAGTCCCAAAATCTCATCGTTGGTGGCGGAACCTTGGGGATGCTCAACTGCC 70
Qy 101 TGCACCTCGCGCGCGGGAATATACCAACGTTACCGTGTGGACCCCTATCTCTGCCCTA 160
Db 71 TCCATCTCGCGCGTGGGGTTACACCAACGTCACCTTCTCGATGTCAATCGCATCCCGT 130
Qy 161 GCGCATCTCGCGGGAACGACGTGAACAAAGTCATTAGCAGTGGCGCAATATTCGAATA 220
Db 131 CACCGATATACGCGCGGCAATGATGAACAACTTGTCTGGCCGACTGTGCACTGCCGATA 190
Qy 221 ACAA-----GACGAATCGAAGTGAATGAGATCTTTGGCGGAAGAGCGGTTTAACGGTT 274
Db 191 GCAAGGTGATGATGAAGACTCAATCTGGAAGCACTTAGCTACGCCGAGCTCAAGGAT 250
Qy 275 GGAAGAAGCAGCCGCTTTTCAAACCGTATATCATGATACGGGCTCTGATGTCGCTT 334
Db 251 GGCTCCACGACCCCTGTCTTCCACCAATCTGCGCAATACAGGCTCTGTGTGGCTGGCT 310
Qy 335 GCTCGCAGGAGGGCTCGATCGCTGGCGGTCCGGTAGCTCGGGGAGGATCCTAATC 394
Db 311 CAACACCAAGTCTATCAAGCAGCTGTAGAAAGATGAGATCGGTGAACATGACCAAGT 370
Qy 395 TGGTGGAACTTACCGCCCGGAGCAATTTCTGTAACTGGCCCGGAAGGGGTGTTGCAAG 454
Db 371 ATACACCTCTCAACACAGCAGAAGATTTCAAGAAAGACCATGCTGAGGGTATCCTGACAG 430
Qy 455 GTGATTTTCGGGTGGAAGGGTACTTTGGCGCTTCGGCGCTGCGTGGGCAATGCAAA 514
Db 431 GTAACTTTCCAGGCTGGAAGGGCTTTTACAAAGCCACGGGTTCTGTGTGGGTTTATGCTC 490
Qy 515 GGAATGCTTAGTGGCAGCAGCAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 574
Db 491 GAAAGCTATGAAGCTGCTTTTGAAGAGAGCAGAGAGCTTGGTGTCAAAATTCATCACTG 550
Qy 575 GCACCCCGCAGGGTGTGTAGTCACTGTTTGAATAATACGATGAAGAGGTGCCG 634
Db 551 GCTCTCCGGAAGAAAGGTGGAGAGTGTATCTTTGAAGAGCGGATGTCGAGGTGCCA 610
Qy 635 TTACGGCGATGGAATAATTTGGAGAGCGGAACGTAATCTCTGTGTGCTGGGCTAGCG 694
Db 611 AGACGGCAGATGTTAAGGAGCAGCAGAGCGGATCGAACTATCTTTCGGCTGGTTCAG 670
Qy 695 CGGTGAGTTCCTAGATTTCAAGATCACTTGCACCAACCGCTTGCACCTGTGTACACA 754
Db 671 CAGAGTTCTTCTCGATTTTGGAGAACAGATCCAGCTACGGCGTGGACCTGGGCCATA 730
Qy 755 TTGGTTTAAACCGGAAGACGTCGTTGTACAAAATATACCGGTTATCTTTAAACATCG 814
Db 731 TCCAGATGACACCGAAGAACCAAGCTGTACAGAACCTGCCACCTCTTTTCAACATCA 790
Qy 815 AACGGGGTTTTTTTGAACCCGATGAGGAGCGCGGTGAGATPAAATAATGCGATGAAC 874
Db 791 ACCAAGGTTTCTTATGGAACCTGATGAGGATCTTCATCAACTCAAGATGTGCGATGAAC 850
Qy 875 ACCGGGCTACACAAATATGTTCCAGAGTCAGACGGCAGCAG-----ATGATGAGCATTCGTT 931
Db 851 ATCCGGGCTATGCACTGGGTTGAAAGCGCTGTTTAAAGTACCCCGAGTCCATCCCT 910
Qy 932 TCGAAAAACCGATTTCCAAAGAACCGCAACCGCGCTTGGGGCTCTGCTGAAGAGA 991
Db 911 TCGCAAGCATCAAGTCCCAACCGAGGCTGAAACGACATGAAGCAGTTTCTGAAGATA 970
Qy 992 CAATGCCCCAGCTGGCAGACGTCCTTCACTTTCAGCTTCGACGCAATTTGCTGGTGGCGATA 1051
Db 971 TCATGCTCAGTTGAGATCGGCGCTTGTTCATGCTCGAATCTGCTGGTGGCTGATA 1030
Qy 1052 CCGGAATCGGAATTTCTGATAGATGACATCGCGAGTACCAGCTTTGTTGTTGGCT 1111
Db 1031 CACAGGATAGAAATTTCTGATCACTTATCATCTCGACATCCCTCACTTTGTATTGCTT 1090
Qy 1112 GTGTGCGAGGAGAGGGTTTAAATACTGCTCTTATTTGGGAATCTCATTTGTCAGC 1171
Db 1091 CAGGTGATTGGGCAAGGGTTACTTTCATATCAATCAATGGAAGTTTCACTCTCACT 1150

Qy 1172 CGATGGAAGTAAAGTCCCGCAAAAAATTCACGAATTAATCAAGTGAACCCGACATTG 1231
Db 1151 GTATGGAGGTACGCTTTGAGGAAAGGTTTCCAAAGTTTCTGGAGATGGCCACGAGAAGT 1210
Qy 1232 CGGCGAACCCGTAACCTGCGGTGATACCTCTGGGGCTTTTGGCGGTCAAATCGTGTGATGG 1291
Db 1211 TTACCGAGTTCTGGGTAAAGATCCTCTGATCGGTTTGGAGCTGACGATGAAGATCATGG 1270
Qy 1292 ATTTTC-----ATGATGTGAAGGAATGGAACCAATGTTCAGTAT 1329
Db 1271 ATTTGCCAAGAGTGTAGAGGATGGACAAATATCAAGAAT 1314

RESULT 14
LOCUS BD173830 1320 bp DNA linear PAT 18-FEB-2003
DEFINITION Compositions for assaying glycoprotein.
ACCESSION BD173830
VERSION BD173830.1 GI:28415163
KEYWORDS WO 02061119-A/4.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1320)
AUTHORS Kozuma,T.; Yoshioka,I.; Arai,M.; Sumitani,J. and Imamu,S.
TITLE Compositions for assaying glycoprotein
JOURNAL Patent: WO 02061119-A 4 08-AUG-2002;
ASAHI KASEI CORP, TAKUJI KOZUMA, ISSEI YOSHIOKA, ISSEI YOSHIOKA, MOTOO ARAI, JUNICHI SUMITANI, SHIGEVUKI IWAMURA
COMMENT OS Artificial Sequence
PN WO 02061119-A/4
PD 08-AUG-2002
PF 30-JAN-2002 WO 2002JP000721
PR 31-JAN-2001 JP 01P 022953,16-FEB-2001 JP 01P 039796 PR
PI 08-AUG-2001 JP 01P 240002
PI TAKUJI KOZUMA, ISSEI YOSHIOKA, MOTOO ARAI, JUNICHI SUMITANI, PI SHIGEVUKI IWAMURA
PC C12Q1/26,C12Q1/37,C12N15/09,G01N33/68,G01N33/72 CC
Compositions for assaying glycoprotein
FH Key Location/Qualifiers
FT CDS (1)..(1320)
FT mutation (1114)..(1115).
FEATURES
source Location/Qualifiers
1..1320
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

Query Match 23.4%; Score 332; DB 6; Length 1320;
Best Local Similarity 55.1%; Pred. No. 6.4e-87;
Matches 719; Conservative 0; Mismatches 570; Indels 15; Gaps 3;

Qy 41 TCACTAAGTCATCATCTCTCTCTGATCGTTGGTGGCGGACTTGGGGCACTCAAGGCTC 100
Db 11 TCACCAAAACAGTCCCAAATTTCTCATCGTTGGCGGAACTTTGGGGATGCTCAACTGCC 70
Qy 101 TGCACCTCGCGCGCGGGGATATACCAACGTTTACCGTGTGGACCCCTATCTCTGCCCTA 160
Db 71 TCACTCTCGCGCGTGGGGTTACCAACGTCATCTGTTCTCGATGTCAATCGCATCCGT 130
Qy 161 GCGCCATCTCCGCGGAAACGAGTCGAACAAAGTCATTAGCAGTGGCGCAATATTCGAATA 220
Db 131 CACCGATATCAGCGCGGCGATGATGTAACAACTTGTCTGGCCGACTGTGCACTGCCGATA 190
Qy 221 ACAA-----GACGAATCGAAGTGAATGAGATCTTTGGCGGAAGAGCGGTTTAAACGGTT 274
Db 191 GCAAGGTGATGATGAAGACTCAATCTGGAAGCACTTAGCTACGCCGAGCTCAAGGAT 250
Qy 275 GGAAGAAGCAGCCGCTTTTCAAACCGTATTAATCATGATACGGGCTCTGATGTCGCTT 334
Db 251 GGCTCCACGACCCCTGTCTTCCACCAATCTGCGCAATACAGGCTCTGTGTGGCTGGCT 310

Sequence split into 13 fragments LOCUS CR382133 Accession CR382133

Qy	335	GCTCGAGGAGGGCCTCGATGCTCGTGGCGTTCGGGTACGTCTCGGCGCAGGATCTCTAATC	394
Db	311	CAACACCAAAAGTCTATCAACGAGCAGCTGGTAGAAGATGAGATCGGTGACCATCGACCAGT	370
Qy	395	TGTTGGAACTTACCCGCGCCGAGCAATTTTCGTAAACTGGCCCGGAAAGGCGTGTTCGAAG	454
Db	371	ATACACCTCTCAACACAGCAGAAGATTTTCAGAAAGACCATGCCCTGAGGGGTATCTTGACAG	430
Qy	455	GTGATTTTTCGGGTTGGAAAGGGTACTTTTCGGCGTTCCTCGGCGTCTGGCTGGGACACATGCAA	514
Db	431	GTAACTTTTCAGGCTCGAAGGGCTTTTACAAGCCACGCGGTCTCTGGTTGCGTTTCATGCTC	490
Qy	515	GGAAATGCTTAGTGGCAGCAGCAGCCGAAAGCAGCGCATGGGTGTAAATTTGTTACTG	574
Db	491	GAAGAAGCTATGAAGAGCTGCTTTTCGAAGAGAGCGAGAGGCTTGGTGTCAAAATTCATC	550
Qy	575	GCACCCCGCAGGGTCGTGTAGTCACTTAATCTTTTGAATAATACGATGTAAAGGTGCGG	634
Db	551	GCTCTCCGGAAGAAAGGTGGAGAGTCTGATCTTTTGAAGACGGCGATGTTTCGAGGTGCCA	610
Qy	635	TTACGGGCGATGGCAAAATTTTGGAGAGCGGAAAGTACATTTCTGTGTGTCTGGGGCTAGCG	694
Db	611	AGACGGCAGATGGTAAGAGCAGACAGAGCGGATCGAACTATTCTTTCGCTGGTCTTCAG	670
Qy	695	CGGCTCAGTTCCTAGATTTCAAGNATCAACTTCGACCAACGGCTTGCAGCCCTGGTACACA	754
Db	671	CAGAGTCTTCTCTCGATTTTGAAGAACAGATCCAGCCTACGGGTGGAGCCCTCGGGCCATA	730
Qy	755	TTGCGTTTAAAAACCGGAAGAACGTCGCTGTGTACAAAAATATACCGGTTATCTTTAAACATCG	814
Db	731	TCCAGATGCACACAGAGAAACCAAGCTGTACAAGAACCTGCCACCTCTTTTCAACATCA	790
Qy	815	AACGGGGTTTTTCTTTGAACCCGATGAGGAGCGCGGTGAGATTAATAATATGGCATGAAC	874
Db	791	ACCAAGGTTTCTTCATGAGAACCTGATGAGGATCTTTCATCAACTCAAGATGTGGGATGAAC	850
Qy	875	ACCGGGCTACACAAATATGTCAGAGTGCAGACGGCAGCG--ATGATGAGCATTTCCGT	931
Db	851	ATCCGGGCTACTGCACCTGGGTGTAAAGCCTGTGTTCTAAGTAGTACCCCGAGTCCATCCCT	910
Qy	932	TCGAAAAAACCCAGATTTCCAAAAGAACGCAAAACGCGCTTCGGGGCCCTGCTGAAAGAGA	991
Db	911	TCGCAAGCATCAAGTGCCAACCGAGGCTGAACGACGCATGAAGCAGTTCCTGAAAGATA	970
Qy	992	CAATGCCCCAGCTGGCAGACCGTCCATTACGCTTCGCACCGATTTGCTGGTGTGCCGATA	1051
Db	971	TCATGCTCTCAGCTTGCAGATCGGCGCTGTGTTCAATGCTCGAACTGCTGGTGGCTGATA	1030
Qy	1052	CCGCGAATCGGGAATTCCTGATAGATCGACATTCGCGAGTACCAAGTCTTGTGTGGGCT	1111
Db	1031	CACAGGATAGAAATGTTCTGATCACTATCAATCTCGACATCCCTCACTTGTCAATGCTT	1090
Qy	1112	GTGTTGGCAGCGGAAGGGGTTTAAATATCTGCCCTTCTATTGGGAAATCTCAITGTTGACG	1171
Db	1091	CAGGTGATTGGGCACGSGTTACGTGCATATCACATCAATTTGGAAGATTCATCTCTGACT	1150
Qy	1172	CGATGGAAGGTAAAGTGGCCGCAAAAATTCAGAAATTAATCAAGTGSAAACCCGACATTCG	1231
Db	1151	GTATGGAGGGTACGCTTGAGGAAAGGTTTGCCAAGTTCTCGAGATGCGCACAGAGAAGT	1210
Qy	1232	CGGCGAACCGGTAACTGCGGCTGATACTCTGGGCGGTTTTTGGCGGTCCAAATCTGCTGATGG	1291
Db	1211	TTACCGAGTTCGGGTAAAGATCTCTGGATCGGTTTGGAGCTGACGATAAGATCATGG	1270
Qy	1292	ATTTTC-----ATGATGTGAAGAAATGGACCAATGTTTCAGTAT	1329
Db	1271	ATTTGCCCAAGAGTGATGAGAGGATGGGACAAATATCAAGAAAT	1314

```
Qy      833 AACCCGATGAGCGGCGGTGAGATTAAATATCGGATGAACACCGCGGCTACACAATA 892
Db      38287 AACCTGATGAAA---CAATGAGATGAAATCTGTAATGAGTTTCCAGGATATACACATT 38231

Qy      893 TGGTCCAGAGTGCAGACGCGCACGATGATGAGCATTCCGTTTCGAAAAAACCCAGATTCCAA 952
Db      38230 ATATAAGTGAACAGAGGAACCAATACTTCAGTTCGGAATATATCAACACTGTTCTCTG 38171

Qy      953 AAGNAGCCGAAACGCGCGTTTCGGGCCCTGCTGAAGAGACAAATGCCCCAGCTGGCAGACC 1012
Db      38170 TTGAAGCTGAAAAAGGAATTAGAAAGCTATTGAAAGAAACTATTCCAGATTTTAGGGATA 38111

Qy      1013 GTCCATTTCAGCTTCGCACGCAATTTGCTGGTGTGCCGATACCGGAATCGCGAATTCCTGA 1072
Db      38110 AACCGATTGTTATGTCCTAAATTTGTTGGTGTACAGATACCCACAGATATTATTATAA 38051

Qy      1073 TAGATCGACATCCGCGAGTACCACAGTCTTGT 1103
Db      38050 TCGACGAATACCCAGAAAGCAATGGTAGTTT 38020
```

Search completed: May 29, 2005, 23:28:26
Job time : 6440 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2005, 19:14:08 ; Search time 837 Seconds
(without alignments)
10035.977 Million cell updates/sec

Title: US-10-622-893A-6
Perfect score: 1419
Sequence: 1 atgggaggttcgggtgacga.....atcacatcatcatcattaa 1419

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1980s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	685	48.3	1314	2	AAV09626
2	685	48.3	1314	2	Az07721 Heat-resi
3	683.4	48.2	1314	2	Aav00732 Aspergill
4	335.2	23.6	1320	2	Aav44893 Fructosyl
5	335.2	23.6	1320	6	Abk90475 Fusarium
6	333.6	23.5	1320	6	Abk90477 Fusarium
7	332	23.4	1320	6	Abk90476 Fusarium
8	332	23.4	1320	6	Abk90478 Fusarium
9	324.4	22.9	1335	13	Adm78788 Fusarium
10	115.2	8.1	1284	13	Adm15772 Fructosyl
11	78.6	5.5	1296	10	Ades1212 Orf20 cod
12	78.6	5.5	72149	10	Ades1173 ML-236B s
13	62.8	4.4	1314	8	Acc48873 Eupenicil
14	57.6	4.1	1314	2	Aax15949 DNA encod
15	53.6	3.8	1314	2	Aat85703 Coding se
16	53.6	3.8	1314	2	Aax15950 DNA encod
17	50.4	3.6	1314	8	Acc48872 Coniochae
18	46.4	3.3	1422	12	Adm78786 Fusarium
19	44.8	3.2	2000	8	Ada71938 Rice gene
20	38.8	2.7	3731	12	Ado39657 Aspergill

21	37.2	2.6	1302	4	AAS56010
22	37	2.6	951	8	ACA37656
23	36.4	2.6	1131	10	ACC61522
24	36.4	2.6	1131	10	ADK64569
25	36.4	2.6	5983	11	ADP65808
26	36.4	2.6	5983	11	ADP65728
27	36.4	2.6	5983	11	ADP65712
28	36.4	2.6	5983	11	ADP65793
29	35.6	2.5	414	9	ACH46271
30	35	2.5	1976	13	ADS56141
c 31	35	2.5	2000	8	ADA71938
32	35	2.5	2953	6	ABV72307
33	34.8	2.5	222	12	ACH82985
34	34.8	2.5	463	4	ABA57449
35	34.8	2.5	463	4	AAI36997
36	34.8	2.5	463	4	AAK31088
37	34.8	2.5	463	4	AAK05487
38	34.8	2.5	463	4	ABS30767
39	34.8	2.5	463	6	ABS05839
40	34.8	2.5	593	12	ACH89285
c 41	34.6	2.4	456	5	AAH66227
c 42	34.6	2.4	456	8	ACA00431
c 43	34.6	2.4	475	11	ADL65880
c 44	34.6	2.4	5059	2	AAH84332
c 45	34.6	2.4	349980	5	AAH68527

ALIGNMENTS

RESULT 1

AAV09626
ID AAV09626 standard; cDNA to mRNA; 1314 BP.

XX AAV09626;

DT 14-MAY-1998 (first entry)

XX A. terreus FAOD-L cDNA.

DE Fructosylamino acid oxidase; FAOD-L; Amadori compound; ds.

KW Fructosylamino acid oxidase; FAOD-L; Amadori compound; ds.

XX Aspergillus terreus.

OS JPI0033180-A.

XX 10-FEB-1998.

PF 24-JUL-1996; 96JP-00194557.

XX 24-JUL-1996; 96JP-00194557.

PR (KYOT-) KYOTO DAIICHI KAGAKU KK.

XX WPI; 1998-172097/16.

DR P-PSDB; AAW39253.

XX Aspergillus terreus recombinant fructosylamino acid oxidase - used for

analysis of "Amadori" compound(s).

PT Example 1; Col 23-26; 18pp; Japanese.

XX This sequence encodes a novel recombinant fructosylamino acid oxidase

(FAOD-L) produced from eukaryotic cells integrated with an RT-PCR product

CC encoding FAOD-L derived from Aspergillus terreus Gp1 (FERM P-15664). FAOD

CC -L can be used for the analysis of "Amadori" compounds

XX Sequence 1314 BP; 294 A; 371 C; 395 G; 254 T; 0 U; 0 Other;

SQ Query Match 48.3%; Score 685; DB 2; Length 1314;

Best Local Similarity 70.7%; Pred. No. 2.5e-218;

Matches 926; Conservative 0; Mismatches 380; Indels 3; Gaps 1;

QY	38	CCGTCTACTAAGTCATCATCTCTCTGATCGTTGGTCGGGACTTGGGGCACTCTCAACGG	97
Db	5	CAGTCACCAAGTCTTTCGTGATATTGATCATCGGGCGGGCACTGGGGTTGCTCAACTG	64
QY	98	CTCTGCACCTCGCGCGCGCGGATATACCAAGCTTACCGTGTGAGACCCCTATCTGTGCC	157
Db	65	CCCTGCTCTTGGCCGAGAGGATACACCAATGTCACTGTCTTACCCCGTACCCCGGTTT	124
QY	158	CTAGCGCATCTCCGCGGAAACGAGTGAACAAAGTCATTAGCAGTGGCCAAATATTCGA	217
Db	125	CATCAGCCATTTCGCGCGGCAACGAGCTCAACAGATCATCTCGTCGGCCAGTACAGCA	184
QY	218	ATAACAAAGACGAATCGAAGTGAATGAGATCTTTGGCGGAAGAGGGTTTTAAACGGTTGA	277
Db	185	GCAAGAGGACGAGTTCGAAGTCAATGAGATTATCGCCGAACAGGCTTCAATGGCTGGA	244
QY	278	AGAACGACCCGCTTTTCAACCGTATTTATCATGATACGGGCCCTGCTGATGTCGTGTGCT	337
Db	245	AAAATGACCCCATCTTCAAGCCGTACTACCACGACACCGCGCTGATGTCCGCCACCA	304
QY	338	CGCAGGAGGSCCTGGATCGCTCGGGTCCGGGTACGTCCGGCGGAGGATCCTAAATCTGG	397
Db	305	CACAGGAAGATTGGAGCGTCTGGGGTCCGGTGGACCTGAAGATGAACCCGATGTAG	364
QY	398	TGGAACCTTACC CGCCGCGACAAATTTCTGTAACCTGCGCCCGAAGCGTGTTCGAAAGTG	457
Db	365	CCGAATTGACTCGCGCGGAGCAGTTCCGCCAGCTGGCCCC---GGGCTCTTGAAGGTA	421
QY	458	ATTTTCGGGTGGAAAGGTACTTTTGGCGGTTCGGCGCTGCTGGGACACATGCAAGGA	517
Db	422	ACTTCCCGGTTGGAGGGGTACCAATTCCTGCTCAAAACGCGGCTGGCGCATCGCGCA	481
QY	518	ATGCCTTAGTGGCAGCAGCAGCGAGCAGCAGCGCATGGGTGTAAATTTGTTACTGGCA	577
Db	482	ACGCCCTTGTGCGCGCGCGGAGGAGCAGCGCCTGGGTGTGCGCTTCGTCCGCGGAT	541
QY	578	CCCGCAGGTCGTGTAGTCACGTTAATCTTTGAAATACGATGTAAAGGTGCGCGTTA	637
Db	542	CGCGCAGGCGAGGTCATCACGTTGATTTTGAACAACGATGTCAGGGTTCGCGTCA	601
QY	638	CGGGCGATGCAAAATTTGAGAGCGGAACGTCATCTCTGTGTGTCTGGGGCTAGCGCGG	697
Db	602	CGCGCGACGCAAGATCTGGCGGCGGAGCAGCATATCTCTCTGCTGTGTGCGCGCGCG	661
QY	698	GTCACTTCTAGATTTTCAACAATCACTTCGACCAACCGCTTCGACCTGTGTACACATTC	757
Db	662	GCCAGTTTCTGGATTTTCAAGGACCACTGCGTCCCACTGCGTGGTCTGCTGCATCC	721
QY	758	CGTTAAACCGGAAGACGTCGCTTGTACAAATAATACCGGTTATCTTTAAACATCGAAC	817
Db	722	AGTTGAAGCCGGAAGAGCGTGCCCGATATTAACAACATGCGCGTGTCTTCAACATCGA	781
QY	818	GGGGGTTTTCTTTGAACCCGATGAGAGCGCGGTGAGATTAAATATGCGATGAACACC	877
Db	782	AGGGGTTCTTCTTCGAGCCGCGATGAGGCGGTGTGAAATCAAGATCTGCGACGCAACC	841
QY	878	CGGCTACACAAATATGTTCCAGAGTCAGACGCGCAGATGATGACATTCCTGTTGAAA	937
Db	842	CCGGGTACACGAATATGACCACGGGGGCCGACGCGCGGTGAGAGATTTCCCTTCGAGA	901
QY	938	AAACCCAGATTCCAAAAGAGCGGAAACGCGCGTTTCGGGCGCTTCGTAAGAGACAAATGC	997
Db	902	AGACGCAAGTTCTTCGAGAGCGGAGATGCGCGTCCGCAAGCTTCTGTCTGAACGATGC	961
QY	998	CCAGCTGGCAGACCGTTCATAGCTTTCGACGCGATTTGCTGGTGTGCCGATACCGCGA	1057
Db	962	CTCAGCTTGGCGACCGCGCGTTAGTTTTCGCAAGGATCTGCTGGTGTGCCGATACCCCA	1021
QY	1058	ATCGCAATTTCTGATAGATCGACATCCGAGTACACACTCTTGTGTTGCGGCTGTGGTG	1117
Db	1022	ATCGGAGTTTATCATTTGACCGTATCCCGAAATACCCGTCGCTGTGTTCTTGGGTGGTG	1081
QY	1118	CGAGCGAAGAGGGTTTTAAATATCTGCTTTCTATTGGGAATCTCATTTGTTGAOCGCGATGG	1177

Db	1082	CTTCAGGACGAGGCTTCAAATATCTTCCCTCGATCGGAAAGCATCATCGCAGACGCCATGG	1141
QY	1178	AAGGTAAGTGGCGCAAAAAAATTCAGAAATTAATCAAGTGGACCCGGACATTCGCGGA	1237
Db	1142	AGGACAAAAACCCCGGCAAAAAATCCACAGCTGATCGCTGGAGCCCGGAAATCGCGATCA	1201
QY	1238	ACCGTAACCTGGCGTGTACTCTGGGGCGTTTTTGGCGGTCCAAATCGTGTGATCGATTTTC	1297
Db	1202	ACCGTAACCTGGGGGACAGATTAGTTCGATTTGGAGGGCCCAACCGGTCATGATTTCA	1261
QY	1298	ATGATGTGAAGGAATGACCAATGTTTCAGTATCGTGTGATTTTCCAAGCT	1346
Db	1262	ATGAAGTGAAGGAGTGACTTAATGTCAACCAAAAGGACATCTCGAAGTT	1310
RESULT 2			
AAZ07721			
ID	AAZ07721	standard; cDNA to mRNA; 1314 BP.	
XX	AAZ07721;		
AC	AAZ07721;		
DT	02-NOV-1999	(first entry)	
DE	Heat-resistant fructosyl amino acid oxidase nucleotide sequence.		
XX	PAOD; thermostable; fructosyl amino acid oxidase; heat-resistant;		
KW	glutathione S-transferase; GST; Amadori compound; ds.		
XX	Aspergillus terreus.		
PH	Location/Qualifiers		
FT	1..1314		
FT	/tag= a		
FT	/product= "fructosyl amino acid oxidase (PAOD)"		
XX	JP11221081-A.		
PN	17-AUG-1999.		
XX	10-FEB-1998; 98JP-00028169.		
PF	10-FEB-1998; 98JP-00028169.		
XX	(KYOT-) KYOTO DALIICHI KAGAKU KK.		
XX	WPI; 1999-520719/44.		
DR	P-PSDB; AAY29100.		
XX	New heat-resistant fructosyl amino acid oxidase - useful for analysis of		
PT	Amadori compounds.		
XX	Claim 1; Page 5-6; 8pp; Japanese.		
PS	This nucleotide sequence encodes an Aspergillus terreus heat-resistant		
CC	(thermostable) fructosyl amino acid oxidase (PAOD). PAOD can be produced		
CC	using standard recombinant methodology. A recombinant fusion protein		
CC	comprising PAOD and glutathione S-transferase (GST) is useful for the		
CC	analysis of Amadori compounds		
XX	Sequence 1314 BP; 294 A; 371 C; 395 G; 254 T; 0 U; 0 Other;		
SQ	Query Match 48.3%; Score 685; DB 2; Length 1314;		
	Best Local Similarity 70.7%; Pred. No. 2.5e-218;		
	Matches 926; Conservative 0; Mismatches 380; Indels 3; Gaps 1;		
QY	38	CCGTCTACTAAGTCATCATCTCTCTGATCGTTGGTCGGGACTTGGGGCACTCTCAACGG	97
Db	5	CAGTCACCAAGTCTTTCGTGATATTGATCATCGGGCGGGCACTGGGGTTGCTCAACTG	64
QY	98	CTCTGCACCTCGCGCGCGGATATACCAAGCTTACCGTGTGAGACCCCTATCTGTGCC	157
Db	65	CCCTGCTCTTGGCCGAGAGGATACACCAATGTCACTGTCTTACCCCGTACCCCGGTTT	124

[illegible]

Db	1142	AGGACAAAACCCCGGCAAAAATCCACAGCTGATCCGCTGGAGCCCGGAAATCCGCGATCA	1201
Qy	1238	ACCGTAACCTGGCTGATATCTCTGGGCGGTTTTGGCGGTCCAAATCGTGTGATGGATTTTC	1297
Db	1202	ACCGTAACCTGGGGGACAGATTAGGTCGATTTGGAGGGCCCAACCGGTCATGGATTCA	1261
Qy	1298	ATGATGTGAAGGAATGGACCAATGTTTCAGTATCGTGATATTTCCAAGCT	1346
Db	1262	ATGAAGTAGGAGTAGGACTTAATGTCCACCAAGGGACATCTCGAAGTT	1310
RESULT 4			
AAV44893			
XX	ID	AAV44893 standard; DNA; 1320 BP.	
XX	AC	AAV44893;	
XX	DT	28-OCT-1998 (first entry)	
XX	DE	Fructosylamine oxidase coding sequence.	
XX	KW	Fructosylamine oxidase; enzyme production; ds.	
XX	OS	Fusarium oxysporum.	
XX	XX	JPL0201473-A.	
XX	PD	04-AUG-1998.	
XX	PF	20-JAN-1997; 97JP-00007101.	
XX	PR	20-JAN-1997; 97JP-00007101.	
XX	PA	(ASAH) ASAH KASEI KOGYO KK.	
XX	DR	WPI; 1998-474488/41.	
XX	DR	P-PSDE; AAW69251.	
PT	Substantially pure microorganism - used for producing fructosylamine oxidase efficiently.		
XX	PS	Claim 1; Page 11-13; 15pp; Japanese.	
CC	This sequence encodes the fructosylamine oxidase produced by the microbe of the invention. The microbe can produce fructosylamine oxidase with high efficiency		
XX	SQ	Sequence 1320 BP; 348 A; 333 C; 330 G; 309 T; 0 U; 0 Other;	
Query Match 23.6%; Score 335.2; DB 2; Length 1320;			
Best Local Similarity 55.3%; Pred. No. 7e-101;			
Matches 721; Conservative 0; Mismatches 568; Indels 15; Gaps 3;			
Qy	41	TCACATAAGTCATCATCTCTCTGATCGTTGGTGC CGGACCTTGGGGACCTCAACGGCTC	100
Db	11	TCACCAACAGTCCCAAATTTCTCATCGTTGGTGGCGGAACTTGGGGATGCTCAACTGCC	70
Qy	101	TGCACCTCGCGCGCGGGATATACCAACGTTACCGTGTGGACCCCTATCTCTGCCCTA	160
Db	71	TCCATCTCGCGCGCTGGGGTTACACCAACGTCATCTGTCTCGATGTCAATCGCATCCCGT	130
Qy	161	GGCGCATCTCGCGCGGAAACGAGTGAACAAAGTCATTAGCAGTGGCCCAATATTCGAATA	220
Db	131	CACCGATATCAGCGCGGCATGATGTAACAAACTTGTGCGCGACTGTCCGATCCCGATA	190
Qy	221	ACAAA-----GACGAATCGAAGTGAATCGAGATCTTGGCGGAAGAGCGCTTTAACCGTT	274
Db	191	GCAAGGTGATGATGAAGACTCAATCTGGAAGAGCACTTAGCTACGCCGACAGCTCAAGAT	250
Qy	275	GGAAGACGACCGCTTTTCAACCGTATTATCATGATACGGCGCTGCTGATGCTGCTT	334
Db	251	GGCTCCACGACCCCTGTCTTCAACCAATTCCTGCCCAATACAGGCTCTGTGCTGGCTGGCT	310

Qy	335	GCTCGCAGGAGGGCCTGGATCGCTCTGGCGTCCGGGTACGTCGGCGCGAGGATCCTTAATC	394
Db	311	CAACACCAAAAGTCTATCAACGACGCTGGTAGAAGATGAGATCGGTGACGACATCGACACAGT	370
Qy	395	TGGTGGAACTTACCCGCCCGGAGCAATTTTCGTAAACTGGCCCCCGAAGGGGTGTTGCAAG	454
Db	371	ATACACCTCTCAACACAGCAGAAGATTTTCAGAAAGACCATGCGCTGAGGGTATCCTGACAG	430
Qy	455	GTGATTTTCGGGTGGAAAGGGTACTTTTGGCGGTTTCGGCGCTGGCTGGGCAACATGCAA	514
Db	431	GTAACCTTTCCAGGCTGGAAGGGCTTTTACAAGCCACGGGTCTCTGGTTGGGTTTCATGCTC	490
Qy	515	GGATGCTCTAGTGGCAGCAGCACGCGNAGCACAGCGCATGGGTGTAATAATTTGTTACTG	574
Db	491	GAAGAAGCTATGAAGAGCTGCTTTTCGAGAGAGCGAGAGGCTTGGTGTCAAAATTCATCACTG	550
Qy	575	GCACCCCGCAGGGTCGTGTAGTCACTGTAATCTTTGAAAAATAACGATGTAAAAAGTGC CG	634
Db	551	GCTCTCCGAGGAAAGGTGGAGAGTCTGATCTTTTGAAGCGGCGATGTTTCGAGGTGCCA	610
Qy	635	TTACGGCGGATGGCAAAATTTTGGAGAGCGGAAAGTAATTCCTGTGTGCTGGGGCTAGCG	694
Db	611	AGACGCAGATGGTAAGGAGCACAGAGCGGATCGAACTAATCTTTCCGCTGGTGTTCAG	670
Qy	695	CGGCTCAGTCTCTAGATTTCAAGAAATCAACTTCGACCACCGCTGACACCTCGGTACACA	754
Db	671	CAGAGTCTTCTCTCGATTTTGAAGAACAGATTCAGGCTTACGCGGTGACACCTCGGCCATA	730
Qy	755	TTGCGTTAAAAACCGAAGAACGTCGCTTGTGTAACAAAAATATACCGGTTTATCTTTAAACATCG	814
Db	731	TCCAGATGACACACAGAAAGAACCAAGCTGTACAAAGAACCTGCCACCTCTTTTCAACATCA	790
Qy	815	AACGGGGGTTTTCTTTGAACCCGATCAGAGCGCGGTGAGATTAAAAATATGCGATGAC	874
Db	791	ACCAAGGTTTTCTCATGGAACCTGATGAGGATCTTTCATCAACTCAAGATGTGCGATGAAC	850
Qy	875	ACCGGGCTACACAAATATGTCGCAGAGTGCAGACGGCAGC --- ATCATGACGATTCGGT	931
Db	851	ATCCGGGCTACTGCACTGGGTTGAAAGCTGTCTTAAGTACCCCGACGTCCATCCCT	910
Qy	932	TCGAAAAAACCCAGATTTCCAAAAGAACCGAAACCGCGCTTCGGGGCCCTGCTGAAAGAGA	991
Db	911	TCGCAAAAGCATCAAGTGCACACCGAGGCTGAACGACGCATGAAGCAGTTCCTGAAGATA	970
Qy	992	CAATGCCCCAGCTGGCAGACGGTCCATTCAGCTTCGCACGCAATTTGCTGGTGTGCGGATA	1051
Db	971	TCATGCTCTACGCTTGCAGATCGGCCGCTGTGTCTATGCTCGAATCTGCTGGTGGCTGATA	1030
Qy	1052	CCGCGAATCGGAAATTCCTGATAGATCGACATCGCAGTAGTACCAAGTCTTGTGTGGGCT	1111
Db	1031	CACAGGATAGAAATGTTCTGTATCACTTATCTCGACATCCCTCACTTGTCAATGCTT	1090
Qy	1112	GTGCTGCGAGCGGAAGGGTTTAAATATCTGCTTCTATTTGGGAATCTCATTTGTCAGC	1171
Db	1091	CAGGTGATTCGGCAGCGGTTTACAAGCATATCACATCAAATTTGAAAAGTTCATCTCTGACT	1150
Qy	1172	CGATGGAAGTAAAGTGCCTCCAAAATATCACGNAATTAATCAAGTGGAAACCCGACACTTG	1231
Db	1151	GTATGGAAGGGTACGCTTGGAGAAAGGTTTGCCAAGTTCTGGAGATGGCGACCGAGAGAT	1210
Qy	1232	CGGCGAACCGTAACTGGCGTGATACTCTGGGGGCTTTTGGCGGTCCAAATCGTGTGATGG	1291
Db	1211	TTACCGAGTCTGGGGTAAAGATCTCTGGATCGTTTGGAGCTGACGATAAGATCATGG	1270
Qy	1292	ATTTTC-----ATGATGTGAAGAAATGGACCAATGTTTCAGTAT	1329
Db	1271	ATTTGCCCAAGAGTGTGATGAGGGATGGAACAAATATCAAGAAT	1314

RESULT 5
ABK90475
ID ABK
XX

AC	ABX90475;
XX	
DT	05-NOV-2002 (first entry)
XX	
DE	Fusarium oxysporum polynucleotide #1.
XX	
KW	Glycoamino acid; glycoprotein; deoxycholic acid; deoxycholic amide; gene;
KW	cholic acid amide octyl glucoside; quaternary ammonium salt; ds;
KW	quaternary ammonium salt cationic surfactant; concanavalin A; betaine;
KW	ascorbic oxidase; 4-(2-hydroxyethyl)-1-piperazinyl group; albumin;
KW	protease; glycoalbumin; bromocresol purple; glycation; diabetes;
KW	ascorbic acid; mutant.
XX	
OS	Fusarium oxysporum.
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	CDS 1..1320
FT	/tag= a
FT	/product= "Fusarium oxysporum polypeptide #1"
FT	/partial
FT	/notes= "No start or stop codon given"
FT	/transl_except= (pos:736..738, aa:Ile)
FT	/transl_except= (pos:1186..1188, aa:Tyr)
XX	
PN	WO200261119-A1.
XX	
PD	08-AUG-2002.
XX	
PF	30-JAN-2002; 2002WO-JP000721.
XX	
PR	31-JAN-2001; 2001JP-00022953.
PR	16-FEB-2001; 2001JP-00039796.
PR	08-AUG-2001; 2001JP-00240002.
XX	
FA	(ASAH) ASahi Kasei Kogyo KK.
XX	
PI	Kouzuma T, Yoshioka I, Arai M, Sumitani J, Imamura S;
XX	
DR	WPI; 2002-599854/64.
DR	P-PSDB; ABG30782.
XX	
PT	Compositions for assaying glycoprotein without interference from globulin
PT	and ascorbic acid in blood, useful in clinical examination including
PT	diagnosis and management of diabetes.
XX	
FS	Claim 12; Page 63-65; 82pp; Japanese.
XX	
CC	The invention relates to compositions for assaying glycoprotein with use
CC	of protease and enzymes acting on glycoamino acids in the presence of at
CC	least 1 of deoxycholic acid or amide, cholic acid amide octyl glucoside,
CC	quaternary ammonium salt, quaternary ammonium salt cationic surfactant,
CC	concanavalin A and betaine and/or ascorbic oxidase and a buffer without a
CC	4-(2-hydroxyethyl)-1-piperazinyl group. The invention also relates to
CC	compositions for assaying albumin containing the protease and enzymes for
CC	determining glycoalbumin and a protein-denaturing agent and/or a compound
CC	having S-S bond and bromocresol purple for evaluation of glycation
CC	proportion of albumin into glycoalbumin with respect to the other
CC	albumins measured separately. The compositions are useful in clinical
CC	examination including diagnosis and management of diabetes. The
CC	compositions can be used in methods which avoid the effects of globulin
CC	and ascorbic acid components and stabilise enzymes such as protease
CC	acting on glycoamino acids. This sequence represents DNA encoding a
CC	polypeptide used in the scope of the invention
XX	
SQ	Sequence 1320 BP; 348 A; 333 C; 330 G; 309 T; 0 U; 0 Other;
	Query Match 23.6%; Score 335.2; DB 6; Length 1320;
	Best Local Similarity 55.3%; Pred. No. 7e-101;
	Matches 72; Conservative 0; Mismatches 568; Indels 15; Gaps 3
QY	41 TCACCTAAGTCATCATCTCTCTCTGATCGTTGGTCCGGGACCTTGGGGACCTCAACGGCTC 100

Db 11 TCACCAACAGTCCCAAAATTCATCGTTGGTGGCGGAACCTTGGGGATGCTCAACTGCC 70
Qy 101 TGCACCTCGCGCCGCGGATATACCAACGTTACCGTGTGACCGCCTATCTGTCCCTA 160
Db 71 TCCATCTCGCGCTCGGGGTACACCAAGTCACTGTTCTCGATGTCATCGCATCCGCT 130
Qy 161 GGGCCATCTCGCGGAAACGAGCTGAA CAAAGTCAATTAGACGTGGCCCAATATTCGAATA 220
Db 131 CACCGATATCAGCGCGGCTGATGTATAA CAAACTTGTGCGCGACTGTGCGCTGCCGATA 190
Qy 221 ACAA-----GACGAAATCGAAGTGAATGAGATCTTGGCGGAGAGGCGTTTAAAGGTT 274
Db 191 GCAAGGTGATGTAAGACTCAATCTGGAAGCACTTAGCTACGCCGAGCTCAAGGAT 250
Qy 275 GGAAGAACACCGCTTTTCAAACCGTATATCATGATACGGGCTGTGATGCTGTGCTT 334
Db 251 GGTCCACGACCTGTCTTCAAACCAATCTGCCCAATACAGGCTGTGCTGGCTGGCT 310
Qy 335 GCTCGAGGAGGCGCTGGATCGCTCGGCGTCCGGTACGTCGGCGGAGGATCCTTAATC 394
Db 311 CAAACCAAAAGTCTATCAAGCAGCTGGTAGAGATCGGTGACGACATCGACCACT 370
Qy 395 TGGTGAACCTACCCCGCGGACCAATTCGTAACTGGCCCCGGAAGGCTGTGCAAG 454
Db 371 ATACACCTCTCAACACAGCAGAGAATTTAGAAAGACCAATGCTGAGGGTATCTCTGACAG 430
Qy 455 GTGATTTTCCGGGTGGAAGAGGTACTTTGCGGTTCCGCGCTGGCTGGGACATGCA 514
Db 431 GTAACTTTCCAGCTGGAAGGCTTTTACAAGCCACGGGTTCTGGTTGGTTTCATGCTC 490
Qy 515 GGAATCCCTTAGTGGCAGCAGCACGCGAAGCACAGCGCATGGGTGTAATAATTTGTTACTG 574
Db 491 GAAAAGCTATGAAAGCTGCTTTTGAAGAGAGCAGAGGCTTGGTGTCAAATTCATCACTG 550
Qy 575 GCACCCCGCAGGTCGTGTAGTCACTTAATCTTTGAAATAACGATGTAAAGGTGCCG 634
Db 551 GCTCTCCGGAAGAAAGGTGGAGTCTGATCTTTGAAGACGCGGATTTGAGGTGCCA 610
Qy 635 TTACGGGCGATGCAAAAATTTGAGAGCGGAAGCTACATTCCTGTGTGCTGGGCTAGCG 694
Db 611 AGACGCGAGATGTAGGAGCACAGAGCGGATCGAACTATCTTTCCGCTGGTGTCTTCAG 670
Qy 695 CGGGTCAGTTCTTAGATTTTCAAGAATCAATCTCGACCAACCGCTTGGACCCCTGGTACACA 754
Db 671 CAGAGTTCTTCTCGATTTTGAAGAACAGATCCAGCTCAGCGCTGGACCCCTGGGCATA 730
Qy 755 TTGCGTTAAACCGGAAGAACGTGGTGTGTAACAAAATATACCGTTATCTTTAAACATCG 814
Db 731 TCCAGATGACACCAAGAAACCAAGCTGTACAAAGAACCTGCGCACCTCTTTTCAACATCA 790
Qy 815 AACGGGGTTTTTTCTTTGAACCGATGAGGAGCGGCTGAGATTAAAATATGCGATGAAC 874
Db 791 ACCAAGTTTCTTCATGGAACCTGATGAGGATCTTCACTCAACTCAAGATGTGGGATGAAC 850
Qy 875 ACCCGGGCTACACAAATATGTTCCAGATGTCAGACGCGCACG---ATGATGAGCATTCGGT 931
Db 851 ATCCGGCTACTGCAACTGGGTTTGAAGAAGCTGTTCTAAGTACCCCGCAGTCCATCCCT 910
Qy 932 TCGAAAAACCCAGATTCGAAAGAGCGAAGCGCGCTTCGGGCCCTGCTGAAAGAGA 991
Db 911 TCGCAAGCATCAAGTGCCAAACCGGCTGAAACGACGATGAAAGCAGTTTCTGAAGATA 970
Qy 992 CAATGCCAGCTGGCAGACCGTCCATTCAGCTTCGCACGCAATTTGCTGGTGTGCCGATA 1051
Db 971 TCAIGCTCAGCTTGAGATCGCGCGCTTGTTCATGCTCGAATCTGCTGGTGGCTGATA 1030
Qy 1052 CCOCGAATCCGAATTCCTGATAGATGACATCCGCAAGTACCAAGCTTTGTGTGGGCT 1111
Db 1031 CACAGGATAGAATGTTCTCTCATCACCTATCATCTCGACATCCCTCACTTGTCAATGCTT 1090
Qy 1112 GTGGTGGCAGCGAAGGGTTTAAATATCTGCTTCTTATGGGAATCTCATTTGTCAGC 1171
Db 1091 CAGGTGATTGCGGACCGGTTTACAAGCATATCATCAATTTGGAAGTTTCATCTCTGACT 1150

Qy 1172 CGATGGAAGGTAAAGTGGCGCAAAAAATTCAGAAATTAATCAAGTGGAAACCGGACATTG 1231
Db 1151 GTATGGAGGTACGCTTGAGGAAAGGTTTGGCCAAGTTCTGGAGATGGCGACCAGAGAAGT 1210
Qy 1232 CGCGAACCGTAACCTGGCGTGATCTCTGGGCGGTTTGGCGGTCCAAATCGTGTGATGG 1291
Db 1211 TTACCGAGTTCTGGGTTAAAGATCCTCTGGATCGGTTTGGAGTGAAGATCAAGATCATGG 1270
Qy 1292 ATTTTC-----ATGATGTGAAGGAATGACCAATGTTTCAGTAT 1329
Db 1271 ATTGGCCCAAGAGTGAATGTAGAGGATGGCAAAATATCAAGAT 1314

RESULT 6
ABK90477

ID ABK90477 standard; DNA; 1320 BP.

XX AC ABK90477;

XX DT 05-NOV-2002 (first entry)

XX DE Fusarium oxysporum polynucleotide #3.

XX KW Glycoamino acid; glycoprotein; deoxycholic acid; deoxycholic amide; gene;
KW cholic acid amide octyl glucoside; quaternary ammonium salt; ds;
KW quaternary ammonium salt cationic surfactant; concanavalin A; betaine;
KW ascorbic oxidase; 4-(2-hydroxyethyl)-1-piperazinyl group; albumin;
KW protease; glycoalbumin; bromocresol purple; glycation; diabetes;
KW ascorbic acid; mutant.

XX OS Fusarium oxysporum.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT CDS 1..1320

XX FT /tag= a /product= "Fusarium oxysporum polypeptide #3"

XX FT /partial

XX FT /note= "No start or stop codon given"

XX FT /transl_except= (pos:736..738, aa:ile)

XX FT /transl_except= (pos:1186..1188, aa:Tyr)

XX FT mutation replace(1115,A)

XX FT /tag= b

XX PN WO200261119-A1.

XX XX 08-AUG-2002.

XX PF 30-JAN-2002; 2002WO-JP000721.

XX XX 31-JAN-2001; 2001JP-00022953.

XX PR 16-FEB-2001; 2001JP-00039796.

XX PR 08-AUG-2001; 2001JP-00240002.

XX XX (ASAH) ASahi KASEI KOGYO KK.

XX PI Kouzuma T, Yoshioka I, Arai M, Sumitani J, Imamura S;

XX XX WPI; 2002-S99854/64.

XX DR P-PSDB; ABG30784.

XX PT Compositions for assaying glycoprotein without interference from globulin
PT and ascorbic acid in blood, useful in clinical examination including
PT diagnosis and management of diabetes.

XX PS Disclosure; Page 69-72; 82pp; Japanese.

XX CC The invention relates to compositions for assaying glycoprotein with use
CC of protease and enzymes acting on glycoamino acids in the presence of at
CC least 1 of deoxycholic acid or amide, cholic acid amide octyl glucoside,
CC quaternary ammonium salt, quaternary ammonium salt cationic surfactant,
CC concanavalin A and betaine and/or ascorbic oxidase and a buffer without a

CC 4-(2-hydroxyethyl)-1-piperazinyl group. The invention also relates to
 CC compositions for assaying albumin containing the protease and enzymes for
 CC determining glycoalbumin and a protein-denaturing agent and/or a compound
 CC having S-S bond and bromocresol purple for evaluation of glycation
 CC proportion of albumin into glycoalbumin with respect to the other
 CC albumins measured separately. The compositions are useful in clinical
 CC examination including diagnosis and management of diabetes. The
 CC compositions can be used in methods which avoid the effects of globulin
 CC and ascorbic acid components and stabilise enzymes such as protease
 CC acting on glycoamino acids. This sequence represents DNA encoding a
 CC polypeptide used in the scope of the invention
 XX
 SQ Sequence 1320 BP; 347 A; 333 C; 330 G; 310 T; 0 U; 0 Other;

Query Match 23.5%; Score 333.6; DB 6; Length 1320;
 Best Local Similarity 55.2%; Pred. No. 2.4e-100;
 Matches 720; Conservative 0; Mismatches 569; Indels 15; Gaps 3;

QY 41 TCACCTAGTATCATCTCTCTCTGATCGTTGGTGGCGGAGCTTGGGACACCTCAACGGCTC 100
 DB 11 TCACCAACAGTCCCAAAATTTCTCATCGTTGGTGGCGGAACTTGGGGATGCTCAACTGCCC 70
 QY 101 TGCACCTCGCGCGCGCGGATATACCAACGTTACCGTGTGGACCCCTATCCTGTCCCTA 160
 DB 71 TCATCTCGCCCGTGGGGTTACCAACGTCACCTGTTCTCGATCGATCGCATCCCGT 130
 QY 161 GCGCCATCTCCGCGGAAACGAGTGAACAAAGTCATTAGCAGTGGCCAAATATTCGAATA 220
 DB 131 CACCGATATCAGCGCGGATGATGTAACAAACTTGTGCGCCGACTGTGCGACTGCCGATA 190
 QY 221 ACAA-----GAGGAATCGAAGTGAATGAGATCTTTGGCGGAGAGCGGTTTAAACGGTT 274
 DB 191 GCAAGGTGATGATGAAGACTCAATCTGGAAGCACTTAGCTAGCGCGCAGCTCAAGGAT 250
 QY 275 GGAAGAACGACCGCTTTTCAAAACGTTATATCATGATACGCGCTCTGCTGATGCTGCTT 334
 DB 251 GGCTCAGACCTCTGCTTCAACCATCTGCCAATACAGGCTCTGTCTGGTGGCT 310
 QY 335 GCTCGCAGGAGGCGCTGGATCGCTGCGGCGTCCGGGTACGTCGCGGCGAGGATCCTAATC 394
 DB 311 CAACACCAAGTCTATCAAGCAGCTGTGTAGAAGATGAGATCGGTGACGACATCGACAGT 370
 QY 395 TGGTGGAACTTACCGCGCGGAGCAATTTGCTAAACCTGGCCCGGAGAGGCGTGTGCAAG 454
 DB 371 ATACACCTCTCAACACAGCAGAGAATTTCAAGAAAGCACATGCGCTGAGGGTATCCTGCAG 430
 QY 455 GTGATTTTCCGGTGTGAAGGGTACTTTTCCGGGTTCCGGGCTGCGTGGGSCACATGSCAA 514
 DB 431 GTAACTTTCCAGGCTGAAGGGCTTTTACAAGCCCAAGGGTTCTGGTGGGTTTCATGCTC 490
 QY 515 GGAATGCTTTAGTGGCAGCAGCAGCGAAGCAGCGCATGCGGTGTAAATTTTGTACTG 574
 DB 491 GAAAGCTATGAAAGCTGCTTTTCGAAGAGAGCGAGAGGCTTGGTGTCAATTCATCACTG 550
 QY 575 GCACCCGCGAGGTCGTGTAGTCAAGTTAATCTTTGAAATAACGATGTAAGAGGTGCGG 634
 DB 551 GCTCTCCGGAAGGAAGGTGGAGAGTCTGATCTTTTGAAGACGCGCATGTTTCGAGGTGCCA 610
 QY 635 TTACGGCGGATGCGAAATTTGGAGAGCGGAAGTCACTTCTGCTGTGGCGGCTAGCG 694
 DB 611 AGACGGCAGATGTTAAGGAGCAGCAGAGCGGATCGAACTATCTTTCCGGTGGTCTCAG 670
 QY 695 CGGTCAGTTTCTAGATTTTCAAGAAATCAACTTTCGACCAACCGCTTGGACCCCTGTGACACA 754
 DB 671 CAGAGTTCTTCTCGATTTTGGAGAACAGATCCAGCTTACGGCTGACCCCTGGGCCATA 730
 QY 755 TTGCGTTTAAACCGGAAGAACGTCGTTGTGTAACAAATATACCGGTTTATCTTTAAACATCG 814
 DB 731 TCCAGATGACACCAAGAGAAACCAAGCTGTACAAGAACCTTGCACCTCTTTTCAACATCA 790
 QY 815 AACGGGGTTTCTTCTTGAACCGGATGAGCGCGGTGAGATTAAATATGCGATCAAC 874
 DB 791 ACCAAGGTTTCTTCAATGGAACCTGATGAGGATCTTTCATCAACTCAAGATGTGGGATGAAC 850

QY 875 ACCCGGGCTACACAAATATGTTCCAGAGTCCAGAGCGCAGC---ATGATGAGCATTTCCGT 931
 DB 851 ATCCGGGCTACTGCTCAACTGGTGTGAAAGAGCGCTGGTCTTAAGTACCCCCAGTCCATCCCT 910
 QY 932 TCCAAAGAAACCCAGATTTCCAAAGAGAGCCGAAAGCGCGCTTTCGGGCGCTTCTGTAAGAGA 991
 DB 911 TCCAAAGCAATCAAGTGCACCGAGGCTGAACGACGATGAAGCAGTTTCTGAAAGATA 970
 QY 992 CAATGCCCGAGCTGGCAGACCGTCCATTTCAGCTTTCGACGCAATTTGCTGGTGTCCGATA 1051
 DB 971 TCATGCTCAGCTTGCAGATCGGCGCTTGTTCATGCTCGAATCTGCTGGTGGCTGATA 1030
 QY 1052 CCGCAATCCGAAATTCCTGATGATGACATCCGAGTACCAGATCTTGTGTGGGCT 1111
 DB 1031 CACAGGATAGAAATGTTCTCTGATCACCCTATCATCTCGACATCCCTCACTTGTCTATTGCTT 1090
 QY 1112 GTGCTGCGAGCGAAGAGGCTTTAAATATCTGCTTCTTATTTGGAAATCTCATTTGTCAGC 1171
 DB 1091 CAGTGAATTCGGACGAGGTTTACATGATATCATCAATTTGGAAAGTTTATCTCTGACT 1150
 QY 1172 CGATGAAGGTAAAGTCCCGCAAAATTCACGAATTAATCAAGTGGAAACCCGACATTTG 1231
 DB 1151 GTATGAGGCTACGCTTGAGGAAGGTTTCCCAAGTTCTGAGATGCGGACGAGAGT 1210
 QY 1232 CGCGCAACCGTAACTGCGGTGATCTCTGCGGCGTTTTGGCGGTCCTCAAAATCGTGTGATGG 1291
 DB 1211 TTACCGAGTTCTGGGTAAAGATCTCTCGATCGGTTGGAGCTGACGATAAGATCATGG 1270
 QY 1292 ATTTTC-----ATGATGTAAGGAATGGACCAATGTTCAAGTAT 1329
 DB 1271 ATTTGCCAAGAGTGTGTAGAGGGATGGACAAATATCAAGAAT 1314

RESULT 7

ABK90476
 ID ABK90476 standard; DNA; 1320 BP.

XX ABK90476;

XX 05-NOV-2002 (first entry)

DE Fusarium oxysporum polynucleotide #2.

XX Glycoamino acid; glycoprotein; deoxycholic acid; deoxycholic amide; gene;
 KW cholic acid amide octyl glucoside; quaternary ammonium salt; ds;
 KW quaternary ammonium salt cationic surfactant; concanavalin A; betaine;
 KW ascorbic oxidase; 4-(2-hydroxyethyl)-1-piperazinyl group; albumin;
 KW protease; glycoalbumin; bromocresol purple; glycates; diabetes;
 KW ascorbic acid; mutant.

XX Fusarium oxysporum.

OS Synthetic.

Key	Location/Qualifiers
CDS	1..1320
FT	/tag= a
FT	/product= "Fusarium oxysporum polypeptide #2"
FT	/partial
FT	/notes= "No start or stop codon given"
FT	/transl_except= (pos:736..738, aa:lle)
FT	/transl_except= (pos:1114..1116, aa:Trp)
FT	/transl_except= (pos:1186..1188, aa:Tyr)
FT	replace(1114,A)
FT	/tag= b
FT	replace(1115,A)
FT	/tag= c

WO200261119-A1.

08-AUG-2002.

30-JAN-2002; 2002WO-JP000721.

XX 31-JAN-2001; 2001JP-0002953.
 PR 16-FEB-2001; 2001JP-00039796.
 PR 08-AUG-2001; 2001JP-00240002.
 PR (ASAH) ASahi KASEI KOGYO KK.
 PA Kouzuma T, Yoshioka I, Arai M, Sumitani J, Imamura S;
 PI WPI: 2002-599854/64.
 XX P-PSDB; ABG30783.
 XX Compositions for assaying glycoprotein without interference from globulin
 PT and ascorbic acid in blood, useful in clinical examination including
 PT diagnosis and management of diabetes.
 XX Example 21; Page 66-69; 82pp; Japanese.
 XX The invention relates to compositions for assaying glycoprotein with use
 CC of protease and enzymes acting on glycoamino acids in the presence of at
 CC least 1 of deoxycholic acid or amide, cholic acid amide octyl glucoside,
 CC quaternary ammonium salt, quaternary ammonium salt cationic surfactant,
 CC concanavalin A and betaine and/or ascorbic oxidase and a buffer without a
 CC 4-(2-hydroxyethyl)-1-piperaziny group. The invention also relates to
 CC compositions for assaying albumin containing the protease and enzymes for
 CC determining glycoalbumin and a protein-denaturing agent and/or a compound
 CC having S-S bond and bromocresol purple for evaluation of glycation
 CC proportion of albumin into glycoalbumin with respect to the other
 CC albumins measured separately. The compositions are useful in clinical
 CC examination including diagnosis and management of diabetes. The
 CC compositions can be used in methods which avoid the effects of globulin
 CC and ascorbic acid components and stabilise enzymes such as protease
 CC acting on glycoamino acids. This sequence represents DNA encoding a
 CC polypeptide used in the scope of the invention
 XX
 SQ Sequence 1320 BP; 346 A; 333 C; 330 G; 311 T; 0 U; 0 Other;
 Query Match 23.4%; Score 332; DB 6; Length 1320;
 Best Local Similarity 55.1%; Pred. No. 8 4e-100;
 Matches 719; Conservative 0; Mismatches 570; Indels 15; Gaps 3;
 41 TCCTAAGTCATCATCTCTCTGATCGTTGGTCCGGGACCTTGGGGACCTCAACGGCTC 100
 11 TCACCAACAGTCCCAAACTCTCATGTTGGTGGCGGAACTTGGGGATGCTCAACTGCC 70
 101 TGCACCTCCGCGCGCGGATATACCAAGTTCACGTTACCGTGTGGACCCCTATCTCTCCCTA 160
 71 TCCATCTCGCCCGTCCGGGGTTACCAACAGTCACTGTCTCGATGTCAATCGCATCCCGT 130
 161 GCGCCATCTCCGCGGAAACGACGTGAACAAAGTCAATAGCAGTGGCCCAATATTCGAATA 220
 131 CACCGATATCAGCCGGGCATGATGTAAACAACTTCTGTGCGGACCTGTGCGACTGCCGATA 190
 221 ACAAA-----GACGAAATCGAAGTGAATGAGATCTTTGGCGGAAGAGCGGTTTAAACGGTT 274
 191 GCAAGGTGATGATGAAGACTCAATCTGGAAGACACTTAGCTACGCGCGAGCTCAAGGAT 250
 275 GGAAGAACGACCCGCTTTTCAACCGGTATATCATGATACGGCCCTGCTGATGCTGCTT 334
 251 GGCTCCACGACCCCTGCTTCCAAACCAATCTGCGCACAAATACAGGCTCTGTGCTGGCTGGCT 310
 335 GCTCGGAGGAGGCGCTTGATTCGCTCGGCGCTCCGGGTACGTCGCGGCGAGGATCCTTAATC 394
 311 CAACACCAAGTCTNTCAACGACGCTGTGTAAGATGAGATCGGTGAGCATCGACCACT 370
 395 TGGTGAACCTTACCCCGCGGACAAATTCGTAAACTGGCCCCCGGAGCGGTGTGCAAG 454
 371 ATACACCTCTCAACACACAGCAGAAGATTTTCAAGAAGACCAATGCTGAGGGTATCCTGACAG 430
 455 GTGATTTTCGGGTGGAAAGGCTACTTTCCGGGTTCGCGGCTTCGGGTGGGCAATGCCAA 514
 431 GTAACCTTCCAGGCTGGAAGGGCTTTTACAAGCCCAACGGGTTCTGTTGGGTTTCATGCTC 490

QY 515 GGAATGCCCTTAGTGGCAGCAGCACGCGAAGCA CAGCGCATGGGTGTAATAATTTTGTACTG 574
 DB 491 GAAAGCTATGAAAGCTGCTTTTCGAAGACGAGAGGCTTGGTGTCAAATTCATCAGCTG 550
 QY 575 GCACCCCGCAGGTCGTGTAGTCACTTAATCTTTGAAAATACGATGTAAAGAGTGGCG 634
 DB 551 GCTCTCTCCGAAGAAAGGTGGAGAGTCTGATCTTTGAAGACGCGCATGTTTCGAGGTGCCA 610
 QY 635 TTACGGCGCATGGCAAAATTTTGGAGAGCGGAACGTACATTCTCTGTGTCTGGGGCTAGCG 694
 DB 611 AGACGGCAGATGTTAGGAGCAGACAGCGGATGCAACTATTCTTCCGCTGGTGTCTCAG 670
 QY 695 CGGGTCAGTTCCTAGATTTTCAAGAATCAACTTCGACCAACCGCTTGGACCCCTGGTACACA 754
 DB 671 CAGAGTTCCTCTCGATTTTGAACACAGATCCAGCTTACGGCTTACGGCTGGACCCCTGGGCATA 730
 QY 755 TTGCGTTAAACCGGAGAAACGTGCGTGTGTACAAAATATACCGGTATCTTTTAAACATCG 814
 DB 731 TCAGATGACACAGAAAGAAACCAAGCTGTACAAGAACCTGCCACCTCTTTTCAAACATCA 790
 QY 815 AACGGGGTCTTTTCTTTTGAACCGGATGAGCGCGGTGAGATTTAAATATATGCGATGAAC 874
 DB 791 ACCAAGTCTTCTCATGGAACCTGATGAGGATCTTCATCACTCAAGATGTGGGATGAAC 850
 QY 875 ACCCGGGCTACAAATATGTTGTCAGAGTGCAGACGGCAGC---ATGATGAGCATTCGGT 931
 DB 851 ATCCGGGCTACTGCAACTGCGTTGAAAGCCTGGTTTAAAGTACCCCGCAGTCCATCCCT 910
 QY 932 TGAATAAACCCAGATTTCCAAAAGAACCGGAGCGGCTTGGGCGCTTCTGCTGAAAGAGA 991
 DB 911 TCGCAAGCATCAAGTGGCCAAACCGGAGCTGAAACGACGATGAAGCAGTTCCTGAAAGATA 970
 QY 992 CAATGCCCGCAGCTGGCAGACCGTCCATTGAGCTTTCGACGCTTCTGCTGTGTGCGGATA 1051
 DB 971 TCAAGCTCAGCTTGCAGATGCGCCCTGTTGTCATGCTGAACTCTGCTGTGTGCGCTGATA 1030
 QY 1052 CCGCGAATCGGAATTCCTGATAGATCGACATCCGCGAGTACCAACAGTCTTGTGTGGGCT 1111
 DB 1031 CACAGGATAGATGTTCTCTGATCACCTATATCTCGACATCCCTCACTTGTCTATTTGCTT 1090
 QY 1112 GTGTCGAGCGGAGAGAGGTTTAAATATCTGCTTCTTATGGGATCTCATTTGTGTGACG 1171
 DB 1091 CAGGTGATTTGCGGACGCGGTTACTTGCATATACATCAATTTGGAAGTTCATCTCTGACT 1150
 QY 1172 CGATGAAAGTAAAGTGGCGCAAAATTCACGAATTAATCAAGTGGAAACCCGCAATTG 1231
 DB 1151 GTATGAGGATGACGCTTGAAGAAAGTTTGCCAAGTTCTGGAGATGGCGACAGAGAT 1210
 QY 1232 CGCGAACCGTAACTGGCGTGATCTCTGGGGCGTTTGGCGGTCCAAATCTGTGTGATGG 1291
 DB 1211 TTACCGAGTTCGGGGTAAAGATCTCTGGATCGGTTTGGAGCTGACGATAAGATCATGG 1270
 QY 1292 ATTTTC-----ATGATGTAAGGATGACCAATGTTCAAGTAT 1329
 DB 1271 ATTTCCCAAGAGTGTATGAGGGATGGCAAAATATCAAGAT 1314

RESULT 8

ABK90478

ID ABK90478 standard; DNA; 1320 BP.

XX

XX ABK90478;

XX AC

XX 05-NOV-2002 (first entry)

DT

XX Fusarium oxysporum polynucleotide #4.

DE

XX Glycino acid; glycoprotein; deoxycholic acid; deoxycholic amide; gene;

XX cholic acid amide octyl glucoside; quaternary ammonium salt; ds;

XX quaternary ammonium salt cationic surfactant; concanavalin A; betaine;

XX ascorbic oxidase; 4-(2-hydroxyethyl)-1-piperaziny group; albumin;

XX protease; glycoalbumin; bromocresol purple; glycation; diabetes;

XX ascorbic acid; mutant.

XX	OS	Fusarium oxysporum.
XX	OS	Synthetic.
XX	FX	
XX	FT	Key
XX	FT	CDS
XX	FT	Location/Qualifiers
XX	FT	1..1320
XX	FT	/product= "Fusarium oxysporum polypeptide #4"
XX	FT	/partial
XX	FT	/notes= "No start or stop codon given"
XX	FT	/transl_except= (pos:736..738, aa:Ile)
XX	FT	/transl_except= (pos:1186..1188, aa:Tyr)
XX	FT	replace(1114,A)
XX	FT	*tag= b
XX	FT	replace(1115,A)
XX	FT	*tag= c
XX	PN	WO200261119-A1.
XX	PD	08-AUG-2002.
XX	PP	30-JAN-2002; 2002WO-JP000721.
XX	PR	31-JAN-2001; 2001JP-00022953.
XX	PR	16-FEB-2001; 2001JP-00037996.
XX	PR	08-AUG-2001; 2001JP-00240002.
XX	PA	(ASAH) ASAHI KASEI KOGYO KK.
XX	PI	Kouzuma T, Yoshioka I, Arai M, Sumitani J, Imamura S;
XX	DR	WPI; 2002-599854/64.
XX	P	P-PSDB; ABG30785.
XX	PT	Compositions for assaying glycoprotein without interference from globulin and ascorbic acid in blood, useful in clinical examination including diagnosis and management of diabetes.
XX	PS	Disclosure; Page 72-75; 82pp; Japanese.
XX	CC	The invention relates to compositions for assaying glycoprotein with use of protease and enzymes acting on glycoamino acids in the presence of at least 1 of deoxycholic acid or amide, cholic acid amide octyl glucoside, quaternary ammonium salt, quaternary ammonium salt cationic surfactant, concanavalin A and betaine and/or ascorbic oxidase and a buffer without a 4-(2-hydroxyethyl)-1-piperazinyl group. The invention also relates to compositions for assaying albumin containing the protease and enzymes for determining glycoalbumin and a protein-denaturing agent and/or a compound having S-S bond and bromocresol purple for evaluation of glycation proportion of albumin into glycoalbumin with respect to the other albumins measured separately. The compositions are useful in clinical examination including diagnosis and management of diabetes. The compositions can be used in methods which avoid the effects of globulin and ascorbic acid components and stabilise enzymes such as protease acting on glycoamino acids. This sequence represents DNA encoding a polypeptide used in the scope of the invention
XX	SQ	Sequence 1320 BP; 346 A; 333 C; 331 G; 310 T; 0 U; 0 Other;
		Query Match 23.4%; Score 332; DB 6; Length 1320;
		Best Local Similarity 55.1%; Pred. No. 8.4e-100;
		Matches 719; Conservative 0; Mismatches 570; Indels 15; Gaps 3;
QY	41	TCACTAAGTCATCATCTCTCCTGATCGTTGGTCCGGGACTTGGGGCACCTCAACGGCTC 100
DB	11	TCACCACAACAGTCCCATAATTCTCATCGTTGGTGGCGGAACCTTGGGGATGTCTCACTGCC 70
QY	101	TGCACCTCGCGGCCCGGATATACCAACGTTACCGTGCTGGACCCCTATCTCTGTCCTTA 160
DB	71	TCCATCTCGCCGTCGGGGTTACACCAACGTCACCTGTTCTCGATGTCAATCGCATCCCGT 130
QY	161	GCCCATCTCCGCGGAAACAGCGTGAAACAAAGTCATTAGCAGTGCGCCAATATTTCGAATA 220

QY 1292 ATTTTC-----ATGATGTGAAGGAATGGACCAATGTTCAAGTAT 1329
 |||||
 Db 1271 ATTTGCCAAGAGTGATGTAGAGGGATGGACAAATATCAAGAT 1314

RESULT 9

ADM78788

ID ADM78788 standard; DNA; 1335 BP.

XX

AC ADM78788;

XX

DT 01-JUL-2004 (first entry)

XX

Fusarium proliferatum fructosylamine oxidase Q2 coding sequence.

XX

KW db; gene; fructosylamine oxidase; fructosylllysine; fructosylvaline;
 Amadori compound; medical examination; diagnosis; serum glucose level;
 food inspection.

XX

OS Fusarium proliferatum.

XX

FH Key Location/Qualifiers

FT CDS

1. .1335

/*tag= a

/product= "fructosylamine oxidase Q2"

FT

FT

XX

PN WO2004029251-A1.

XX

PF 16-SEP-2003; 2003WO-JP011766.

XX

XX

PR 24-SEP-2002; 2002JP-0027214.

PR

24-OCT-2002; 2002JP-00309734.

XX

XX

PA (ARKR-) ARKRAY INC.

XX

XX Yoshida N, Tani Y, Yonehara S;

PI

XX WPI; 2004-316127/29.

DR

P-PSDB; ADM78789.

XX

Fusarium proliferatum-originated fructosylamine oxidase useful in

PT measurement of Amadori compound and particularly applicable in medical

PT examination including diagnosis and management of diabetes.

XX

PS Claim 8; SEQ ID NO 5; 50pp; Japanese.

XX

CC The invention relates to a novel fructosylamine oxidase isolated from
 CC Fusarium proliferatum. The enzyme has the following physicochemical
 CC characteristics: (a) activity against fructosylllysine equivalent to or
 CC higher than that against fructosylvaline; (b) optimum pH for enzyme
 CC activity of 7.5; (c) enzyme stability optimally at 30-40 deg C; and (d) a
 CC molecular weight of about 39 kDa as measured by SDS-PAGE and about 39.4
 CC kDa by gel filtration. Also disclosed is a similar Fructosylamine oxidase
 CC characterized by: (i) activity against fructosylvaline and no detectable
 CC activity against fructosylllysine; (ii) optimum pH for enzyme activity of
 CC 7; (iii) enzyme stability optimally at 30-40 deg C; and (iv) a molecular
 CC weight of about 49 kDa as measured by SDS-PAGE and about 58 kDa by gel
 CC filtration. The oxidases are useful in the measurement of Amadori
 CC compounds and particularly applicable in medical examination including
 CC diagnosis and management of diabetes by controlling serum glucose level
 CC for treatment and prevention of complications, and in food inspection.
 CC This sequence represents the coding sequence for the fructosylamine
 CC oxidase designated Q2.

XX

SQ Sequence 1335 BP; 363 A; 328 C; 333 G; 311 T; 0 U; 0 Other;

XX

Query Match

Best Local Similarity 22.9%; Score 324.4; DB 12; Length 1335;

XX

Matches 696; Conservative 0; Mismatches 571; Indels 9; Gaps 2;

XX

QY 31 GCTCTCGCGCTCACTAAGTCAATCATCTCTCCTGATCGTTGGTGCCGGGACTTGGGGACC 90
 Db |||||
 QY 13 GTTCCCCCGCTCAATAAGGACTCAGGGAATCTCATCGTTGGTGGCGAACTTGGGATGC 72
 Db |||||
 QY 91 TCNAAGGCTCTGCACCTCGCGCGCGCGGATATACCAAGTTACCTGCTGGAGCCCTAT 150
 Db |||||
 QY 73 TCAACTGSCCTCCATCTCGCCCGTCTGGGGGTACACCAACGTCACTGTTCTCGATGTCAT 132
 Db |||||
 QY 151 CTTGTCCTTAGCGCCATCTCCGCGGAAACGAGCTGAAACAAAGTCATTAGCAGTGGCCAA 210
 Db |||||
 QY 133 CGATCCCGTACCAGATATACCGCGGCGATGATGTAACAAACTTCTTAACAGACTAGGC 192
 Db |||||
 QY 211 TATTCGAATAACAAAGACGAATTCGAAG-----TGAATGAGATCTTGGCGGAAGAGCG 264
 Db |||||
 QY 193 ACTTCTGATAGTAAAGGCGATGACGAAGACTCAATCTGGAAGCTCTTACGTACGCGCA 252
 Db |||||
 QY 265 TTTAAGGTTGMAAGAACGACCGCTTTTCAAAACGTTATATCATGATACGGCGCTGCTG 324
 Db |||||
 QY 253 GCTCAAGGATGGCTCCATGATCCCATCTTCCAACTTTCGCCACAATACAGGAGCTGTC 312
 Db |||||
 QY 325 ATGTCCTGCTTGTCTCGCAGGAGGSCCTGGATCGCTCGGGCTCGGGTACGTCCTCGGGCGAG 384
 Db |||||
 QY 313 ATGGCTGGCTCAACACAAATCTATCAAGAGCTGTGTAGAGATGAGATCGGTGAGCAC 372
 Db |||||
 QY 385 GATCTTAATCTGCTGGAACCTTACCCCGCCGGAGCAATTTGCTAAACTGGCCCCGGAAGGC 444
 Db |||||
 QY 373 ATCGACCACTATACACCTCTCAACACAGCAGAAGAATTCAGAAGGACTATGCGCGAGCT 432
 Db |||||
 QY 445 GTGTTCAAGGTGATTTTTCGGGTTCGAAAGGTACTTTTCGGCGTTCGGCGCTGGCTGG 504
 Db |||||
 QY 433 ATTCTGACAGGTGATTTTCTAGGCTGGNAGGGCTTTTACAAGCCCAAGAGGTTCAAGTGG 492
 Db |||||
 QY 505 GCACATGCAAGGAATGCTTTAGTGGCAGCAGCAACGCGAAGCAGCAGCGCATGGGTGTAATA 564
 Db |||||
 QY 493 GTTCAITGCCAAGAAAGGCTATGAAGCTGCTTTGAAGAGAGCCAGAGACTTGGTGTCAAG 552
 Db |||||
 QY 565 TTTGTTACTGGCACCCCGCAGGCTGCTAGTACCGTTAATCTTTGAAATAACGATGTA 624
 Db |||||
 QY 553 TTCATCACTGGCTCTCCGGAAGCAAGGTCGAGAGTCTGGTCTTTGAAAGCTGGTGATGTC 612
 Db |||||
 QY 625 AAAGGTGCGGTTACGGGCGATGSCAAANTTTGGAGAGCGGAACGTATCTCTGTGTGCT 684
 Db |||||
 QY 613 AAAGGTGCAAAAACACAGCATGGAAGGAACACAGAGCGGATCGAACAAATTTCTTCGGCT 672
 Db |||||
 QY 685 GGGGCTAGCGCGGTCAGTCTCTAGATTTCAAGAACTCAACTTCGACCAACCGCTTGAGCC 744
 Db |||||
 QY 673 GGTGCTCTCAGCAGAGTCTCTCCCTCGATTTTGAGAACAGATCCGCTCTACGGATGACT 732
 Db |||||
 QY 745 CTGGTACACATTGCGTTAAACCCGGAAGAACGTCGCTGTGTACAAAATAATACCGGTTATC 804
 Db |||||
 QY 733 CTGGGCGCATATCCAGATGACAGCAGAGGAAACAAAGCTGTACAAGGAACCTTCCCGCCCTT 792
 Db |||||
 QY 805 TTTAATCATGAACGGGGTTTTTCTTTTGAACCGGATGAGAGCGCGGTGAGATTAATAATA 864
 Db |||||
 QY 793 TTTCAATATCAACACAGGGCTCTCTCATGGAACCCGATGAGACTTGCATCAACTCAAGATG 852
 Db |||||
 QY 865 TGGATGAAACACC CGGCTCTACAAATATGCTCCAGAGTGCAGACGCA---CGATGATG 921
 Db |||||
 QY 853 TGGATGAAACATCCCGGATCTGCAATTTGGTTGACAAACCTGGTTCCAAATACCCCGCAG 912
 Db |||||
 QY 922 AGCATTTCCGTTGAAAAAACCCAGATTTCAAAAAGAGCCGAAACGCGCGTTCGGGGCCGTG 981
 Db |||||
 QY 913 TCCATCCCTTCGCAAAAGTATCAAGTGCCTCAATTTGAGGCTGAAACGAGCATGAAGCAATTT 972
 Db |||||
 QY 982 CTGAAGAGCAATGCCCCAGCTGGCAGACCGTCCATTAGCTTCGACCGCATTTGCTGG 1041
 Db |||||
 QY 973 CTGAAGAGACATCATGCTCTCAGCTCGCAGATCGGCCACTTGTTCATGCTCGAATCTGCTGG 1032
 Db |||||
 QY 1042 TGTGCCGATACCGCGAATTCGGAATTTCTGTAGATCGACATCGCAGTAGTACCAAGTCTT 1101
 Db |||||
 QY 1033 TGGCCGATACACAGATAGATGTTTCTGATCAGTATCACCCTCGACACCCATCGCTT 1092
 Db |||||
 QY 1102 GTGTTGGGCTGTGTCGCGAGCGGAAGGGTTTTAAATATCTGCTTCTATTGGGAATCTC 1161
 Db |||||


```
Db 1959 GCCGGTGGGCGAGCTGCTGGTACAGCTATCGAGCGGATGCGAGCGCTTCTCAGGAGAA 1900
Qy 556 GGTGTAAATTTGTACTGGCACCAGGTCGTGTAGTCAGCTTAATCTTTGAAAT 615
Db 1899 GGGGTGAATACATTTCCGGGAAGACGAAGGGGATGTGCAAGAGTTGATACCTGGATCAT 1840
Qy 616 A-----ACGATGTAAAGGTGCGCTTACGGGCGATGCGAAATTTGGAGAGCGGACGT 669
Db 1839 ATTACGGGCGAGTCAAGGAGTGTCTCAGACAGATGCGAGACACGAGCGCGGATGTG 1780
Qy 670 ACATTCCTGTGTGCTGGGGCTAGCGCGGTGAGTTCCTAGATTTCAAGAATCACTTCGA 729
Db 1779 GTCATCTTAGCGCAGAGCTTGGACACCTTCGCTACTAGAGTGAAGGGCCAGTTAACT 1720
Qy 730 CCAACCCCTTGGACCCCTGGTACATTCGCTTAAACCGGAAGAACCGCTGCTTGTACAA 789
Db 1719 GCCAAGGGGACAGTGTGCTCATATACAGTGAACCCCTTCCGAAGCAACATATGCCC 1660
Qy 790 AATATACGGTTATCTTTAAACATCGAACCGGGGTTTTTCTTTGAACCCGATGAGGAGCGC 849
Db 1659 TCGATGCCATATGGATATCTCGAGTGGGCTACTTCTCCACCTCAAGAAAGATGGG 1600
Qy 850 GGTGAGATTAAATATCGGATGAACACCCGGGCTACAAATATGTCAGAGTGCAGAC 909
Db 1599 ATCTTCAAGATGGCTACAGTCAATTCATCACCACCGTGCACACCAAGTCTGGTATC 1540
Qy 910 GGCACGATGATGAGCATTCGTTTGGAAACCAACGAGA-----TTCCAAAGAAAGCC 960
Db 1539 ACCATCTCTGTCCACACAGCTTTGTTTCAGGACCCACCCAGTGGCCCTCTTTCAGATC 1480
Qy 961 GAAACGGCGTTCGGGCGCTGCTGAAGAGACATGCCCGAGTGGCAGACCGTCCATTTC 1020
Db 1479 GAAGCCCAATGCGCCGGAATCTCGAGCTGTGCTCCCGAATTTGCAGACAGACCCCTTC 1420
Qy 1021 AGCTTCGACGCAATTTGCTGTGTGCGGATACCGCAATCGCGAATTCCTCATAGATCGA 1080
Db 1419 TGCTATACAGCGCTGTCTGGGATGCTGATAGGCTGATCGCATTTCTTGTGACCCCT 1360
Qy 1081 CATCCGAGTACACAGTCTTGTGTTGGGCTGTGTCGAGCGGAGCGGAAGGGTTTAAATAT 1140
Db 1359 CATCTGCTCACAAAAGTCTCTCTCGCAACCGCGCGATCAGCACATGGGTTCAAGTTC 1300
Qy 1141 CTGCTTCTATTGGGAATCTCATTTGACGCGATGAAGTAAAGTCCGCGCAAAAT 1199
Db 1299 CTTCTGCTGTGAAATATGTGCGGATTTGCTTGAGGGGACTCTGGACCCAGAGAT 1241
```

RESULT 13

ACC48873

ID ACC48873 standard; cDNA; 1314 BP.

XX AC ACC48873;

XX 11-AUG-2003 (first entry)

XX DE Eupenicillium terrenum fructosyl peptide oxidase coding sequence.

XX KW Fructosyl peptide oxidase; enzyme; diabetes; diagnosis; gene; ss.

XX OS Eupenicillium terrenum.

Key	Location/Qualifiers
FT CDS	1..1314
FT	/*tag= a
FT	/product= "Fructosyl peptide oxidase"

XX EP1291416-A1.

XX PD 12-MAR-2003.

XX PF 03-SEP-2002; 2002EP-00256108.

XX PR 04-SEP-2001; 2001JP-00266665.

PR 12-DEC-2001; 2001JP-00378151.
PR 06-AUG-2002; 2002JP-00228727.

XX PA (KIKK) KIKKOMAN CORP.

XX PI Kurosawa K, Hirokawa K, Kajiyama N;

XX WPI; 2003-395517/38.
XX DR P-PSDB; ABR41990.XX PT New fructosyl peptide oxidases useful as enzymes for clinical diagnosis.
XX PS Claim 22; Page 34-36; 59pp; English.

XX CC The present sequence is the coding sequence of Eupenicillium terrenum ATCC 18547 cDNA encoding a novel fructosyl peptide oxidase (FPO). The cDNA was obtained by PCR and RACE amplification using primers based on peptide fragments of the purified enzyme. It is deposited in plasmid puc-EFP as FERM BP-8131. FPO enzymes of the invention act on fructosyl valyl histidine in the presence of oxygen and catalyse a reaction that produces alpha-ketoaldehyde, valyl histidine and hydrogen peroxide. The E. terrenum FPO exhibits 182% relative activity for fructosyl glycine and 9.78% for epsilon fructosyl lysine when compared to 100% activity for fructosyl valyl histidine. It shows optimal activity in the pH range 6.0-8.0 (preferably pH 7.0) and at temperatures around 35 degrees C, and is stable in the pH range 5.0-9.0 and at temperatures up to about 45 degrees C. The Km value for fructosyl valyl histidine is 4.25 mM. The E. terrenum cDNA can be used for mass production of the enzyme by recombinant methods. Stable FPO enzymes such as E. terrenum FPO are useful in clinical diagnosis, e.g. of diabetes

XX SQ Sequence 1314 BP; 328 A; 309 C; 350 G; 327 T; 0 U; 0 Other;

Query Match	4.4%	Score 62.8;	DB 8;	Length 1314;
Best Local Similarity	48.2%;	Pred. No. 2e-09;		
Matches	298;	Conservative 0;	Mismatches 287;	Indels 33; Gaps 3;

Qy	674	TCCTGTGTCTGGGGCTAGCGCGGTGCTCTCTAGATTTTCAGAAATCAACTTCGACCAA	733
Db	650	TCTTGGCTGTGTGGCTGGAGTCCACCTTGGTGGATCTAGAAGATCAGTGTGTTTCAA	709
Qy	734	CCGCTTGGACCCCTGGTACACATTCGCGTTAAACCGGAAGAACGTCGTTGTACAAAATA	793
Db	710	AGCGCTGGGTTTTCGCTCATATTCAACTCACCCAAAGAACGGACGCGTACAAAGATG	769
Qy	794	TACCGGTATCTTTAAACATCGAACCGGGGTTTTTCTTTGAACCGGATGAGGAGCGCGTG	853
Db	770	TGCTGTGTCTATGATGGTGAATATGGGTTCTTTTGGAGCCCAACGAGTA---TGGGG	826
Qy	854	AGATTAAATATCGGATGAACACCCGGGCTACACAA-----ATATGGTCC	898
Db	827	TGATCAAAAGTCTGTGACGAGTTCCTCTGGTTTCTCTCGCTTCAAACTGCATCAACCGTACG	886
Qy	899	AGAGTCGACAGCGGCACGATGATGAGCATTCGTTTCGAAAAAACCCAGATTCCTCAAAAGAG	958
Db	887	GGGCTGCATCTCCCAAGATGATATCCGTACCGGATCACACGCCCAAGCATCCCAACAGATA	946
Qy	959	-----CCGAAACCGCGGTTCCGGGCCCTGCTGAAAGAGACAAATGCCCCAGC	1003
Db	947	CCTACCTGTATGCTCCGAGTCCACATACGCAAGCGATCGCAAGTTCCTGCGAGAT	1006
Qy	1004	TGGACAGACCGTCAATTCAGCTTCGACGCAATTTGCTGGTGTGCGGATACCGCGGAATCGCG	1063
Db	1007	TTAAAGACAAAGGAGCTCTTCAACCGTACCATGTGCTGTGTGTACAGATACGCGCGCATGCTA	1066
Qy	1064	AATTCTCTGATAGATCGACATCCGACGATCCACAGTCTTGTTGGGCTGTGGTCCGAGCG	1123
Db	1067	ACTTATGATTTGCGAACACCCGAAAGTGAAGAAATTTTCATTCTGCGCCACTGGAGATAGCG	1126
Qy	1124	GAAGAGGGTTTAAATATCTGCCTTCTATTGGGAATCTCATTTGTGACCGCATGGAAGGTA	1183
Db	1127	GACATTCCTCAAGCTGTTGCCAAACATCGGGAAACACGTTGTTAGCTTTTAGAGGGAT	1186

```
QY 1184 AAGTCCGCAAAAATTCAGAAATTAATCAAGTGGAAACCGGACATTTGGCGGAACCGTA 1243
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1187 CTCATCGCAGGAATGGCTGGTCTGGAGATGGAGACCGGAGGTGATGCTCTTAGAT 1246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1244 ACTGGCGTGATCTCTGG 1261
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1247 CTAGACGCGGTGCTCCGG 1264
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
AA15949
ID AAX15949 standard; DNA; 1314 BP.
XX
AC AAX15949;
XX
DT 14-MAY-1999 (first entry)
XX
KW DNA encoding a fructosylamino acid oxidase.
DE
XX Fructosylamino acid oxidase; FAOD-P; Penicillium janthinellum S-3413; ds.
XX
OS Synthetic.
OS Penicillium janthinellum.
XX
FH Key Location/Qualifiers
FT 1..1314
FT CDS /*tag= a
XX
PN JP11046769-A.
XX
PD 23-FEB-1999.
XX
PF 05-AUG-1997; 97JP-00210609.
XX
PR 05-AUG-1997; 97JP-00210609.
XX
PA (KYOT-) KYOTO DAIICHI KAGAKU KK.
XX
WPI; 1999-208112/18.
DR P-PSDB; AAW97386.
DR
PT New DNA coding fructosylamino acid oxidase - synthetically designed and
PT recombinantly prepared.
XX
PS Claim 1; Page 8-10; 21pp; Japanese.
XX
CC The present DNA sequence encodes a fructosylamino acid oxidase (FAOD-P).
CC The present sequence is a synthetic sequence, derived from the wild type
CC FAOD-P sequence of Penicillium janthinellum S-3413. The sequence was
CC constructed using primers AAX15951-6006
XX
SQ Sequence 1314 BP; 408 A; 168 C; 244 G; 494 T; 0 U; 0 Other;

Query Match 4.1%; Score 57.6; DB 2; Length 1314;
Best Local Similarity 54.3%; Pred. No. 1.1e-07;
Matches 139; Conservative 0; Mismatches 114; Indels 3; Gaps 1;

QY 628 GTGCGGTACGGCGATGCAAAATTTGGAGAGCGGAACGTACATTCCTGTGTCTGGG 687
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 604 GGTGTTGAAACTGTGTGATGCTACTAATAATATTTTGCTGATAAAGTTGTTTAGCTGCTGGT 663
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 688 GCTAGCGCGGTTCAGTTCTCTAGATTTCAAGAACTCAACTTCGACCAACCGCTTGGACCTG 747
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 664 GCTTGGCTTCTACTTTAGTTGATTTAGAGATCAATGTGTTTCTAAGGCTTGGGTTTTT 723
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 748 GTACACATTCGGTTAAACCGGAAGACGTGCGTTGTACAAAATATACCGTTATCTTT 807
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 724 GCTCATATACAATTAACCTCCACAGAATCTGCTCAATATATAAAGATGTTCCAGTTGTTAT 763
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 808 AACATCGAACGGGGTTTTTCTTTGACCCGNTGAGGAGCGCGGTGAGATTAAATATGC 867
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 784 GATGGTGATTATGGTTTTTTTTTTTGAACCAAAAT---GAACATGGTGTATAAAGTTGT 840
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 868 GATGAACACCCGGGCT 883
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 841 GATGAATTTCCAGGTT 856
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
AAT85703
ID AAT85703 standard; cDNA to mRNA; 1314 BP.
XX
AC AAT85703;
XX
DT 16-FEB-1998 (first entry)
XX
DE Coding sequence for fructosyl amino acid oxidase.
XX
KW Fructosyl amino acid oxidase; FAOD-P; enzyme; amadori compound;
KW alpha-ketoaldehyde amine derivative; amadori detection; ss.
XX
OS Penicillium janthinellum.
XX
FH Key Location/Qualifiers
FT 1..1314
FT CDS /*tag= a
XX
PN WO9721818-A1.
XX
PD 19-JUN-1997.
XX
PF 13-DEC-1996; 96WO-JP003651.
XX
PR 14-DEC-1995; 95JP-00325672.
XX
PA (KYOT-) KYOTO DAIICHI KAGAKU CO LTD.
XX
PI Kato N, Sakai Y, Tani Y, Fukuya H;
XX
WPI; 1997-332789/30.
DR P-PSDB; AAW24134.
DR
PT Fructosyl amino acid oxidase which oxidises amadori compounds and related
PT DNA - used for detection of amadori compounds in foods and in body fluids
PT such as blood.
XX
PS Claim 1; Page 33-36; 53pp; Japanese.
XX
CC This sequence represents the coding sequence for the fructosyl amino acid
CC oxidase (FAOD-P) of the invention. This sequence was isolated from
CC Penicillium janthinellum S-3413 (FERM BP-5475). FAOD-P oxidises amadori
CC compounds in the presence of oxygen, producing alpha-ketoaldehyde, amine
CC derivatives and hydrogen peroxide. FAOD-P is used in the detection of
CC amadori compounds in, e.g. foods such as soy sauce, and in body fluids
CC such as blood. This sequence, vectors containing it and host cells
CC transformed by the vector are all used for the recombinant production of
CC FAOD-P
XX
SQ Sequence 1314 BP; 339 A; 323 C; 345 G; 307 T; 0 U; 0 Other;

Query Match 3.8%; Score 53.6; DB 2; Length 1314;
Best Local Similarity 53.1%; Pred. No. 2.5e-06;
Matches 138; Conservative 0; Mismatches 119; Indels 3; Gaps 1;

QY 628 GTGCGGTACGGCGATGCAAAATTTGGAGAGCGGAACGTACATTCCTGTGTCTGGG 687
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 604 GGTGTTGAGACATAGATGGAAACAAATACTTCGCCGACAAGGTGTTTGGCGGCTGGT 663
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 688 GCTAGCGCGGTTCAGTTCTCTAGATTTCAAGAACTCAACTTCGACCAACCGCTTGGACCTG 747
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 664 GCTTGGAGTTCGACGTGTAGTAGATTTGGAGGACCAATGTGTTTTCGAAGGCTGGGCTTC 723
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 748 GTACACATTCGGTTAAACCGGAAGACGTGCGTTGTACAAAATATACCGTTATCTTT 807
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 724 GCTCATATCAACTCAGCCCCCAAGAAATCGGCCCATGACAGACGTCGCCGTAGTATAC 783
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Qy	808	AACATCGAACGGGGTTTTTTCTTTTGAACCCGATGAGGAGCGCGTGAGATTAAATATGC	867
Db	784	GACGGTATTATGCTTTTCTTCGAGCCC---AACGACACGGAGTATCAAGTCTGC	840
Qy	868	GATGAACACCCGGGCTACAC	887
Db	841	GATGAGTTCCCCGGGTTCTC	860

Search completed: May 29, 2005, 20:15:46
 Job time : 844 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2005, 19:21:38 ; Search time 5114 Seconds
(without alignments)
10561.830 Million cell updates/sec

Title: US-10-622-893A-6
Perfect score: 1419
Sequence: 1 atggggggttcgggtgacga.....atccatcatcatcattaa 1419

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_ges1.*
9: gb_ges2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	237.4	16.7	782	7	CO142686
2	118.4	8.3	1040	9	CNS07A13
3	99.2	7.0	906	9	CNS07AGE
4	83.4	5.9	440	1	AA787493
5	72.6	5.1	397	4	BI187569
6	66.2	4.7	677	7	CN241854
7	65.2	4.6	414	4	BI187570
8	64.6	4.6	599	1	AJ636269
9	64.6	4.6	628	4	BM869701
10	63.4	4.5	459	7	CF188389
11	57	4.0	422	4	EG465351
12	55.2	3.9	758	7	CF871025
13	55.2	3.9	816	6	CB901239
14	55.2	3.9	822	6	CB901049
15	55.2	3.9	822	7	CF870843
16	55	3.9	613	7	CK907858
17	53.8	3.8	491	7	CF188390
18	50.2	3.5	878	9	CNS07990
19	50	3.5	795	7	CF684496
20	49.6	3.5	498	1	AA901536
21	45.8	3.2	686	8	AQ447867
22	44.6	3.1	375	1	AJ237240
23	43.4	3.1	686	7	CF883318
24	43.4	3.1	756	6	CB909424

25	43	3.0	486	7	CK909112
26	42.6	3.0	746	7	CF684508
27	42.6	3.0	749	7	CF718110
C 28	40.8	2.9	794	9	CC997434
C 29	40.6	2.9	1009	9	CC782545
30	40.4	2.8	934	9	CNS07BF2
31	40	2.8	551	7	CO056263
32	40	2.8	598	7	CO053986
C 33	39	2.7	683	6	CD309225
C 34	38.4	2.7	438	6	CA962243
35	38.4	2.7	468	7	CK731051
36	38.2	2.7	349	5	BQ497928
C 37	38.2	2.7	1255	8	CC309899
38	38	2.7	337	2	AW309033
39	38	2.7	431	6	CB025818
40	37.6	2.6	606	7	CF573978
41	37.6	2.6	833	9	CC838405
42	37.4	2.6	1101	9	CNS00370
C 43	37.2	2.6	436	6	CD734891
C 44	37.2	2.6	525	1	AA029863
C 45	37.2	2.6	584	5	BX275202

ALIGNMENTS

RESULT 1
CO142686
LOCUS
DEFINITION
EST837357 Aspergillus flavus Normalized cDNA Expression Library
Aspergillus flavus cDNA clone NAFES50 5' end similar to (O42629)
Fructosyl amine:oxygen oxidoreductase, mRNA sequence.
CO142686
CO142686.1 GI:48896687
EST.
Aspergillus flavus
Aspergillus flavus
Aspergillus flavus
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
1 (bases 1 to 782)
Yu, J., Whitelaw, C.A., Nierman, W.C., Bhatnagar, D. and Cleveland, T.E.
Aspergillus flavus expressed sequence tags for identification of
genes with putative roles in aflatoxin contamination of crops
FEMS Microbiol. Lett. (2004) In press
Contact: Yu J
Food and Feed Safety Research Unit
USDA/ARS, Southern Regional Research Center
1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA
Tel: 504 286 4405
Fax: 504 286 4419
Email: jiyu@arrrc.ars.usda.gov
Contact Dr. Yu at USDA/ARS SRRR (jiyu@arrrc.ars.usda.gov) for clone
information
PCR Primers
FORWARD: M13P
BACKWARD: M13R
Seq primer: M13 Forward
POLYA=No.

FEATURES
source
Location/Qualifiers
1..782
/organism="Aspergillus flavus"
/mol_type="mRNA"
/strain="NRRL 3357"
/db_xref="taxon:5059"
/clone="NAFES50"
/sex="asexual mycelia"
/cell_type="mycelia"
/dev_stage="developmental stages from 18 to 96 hours"
/lab_host="E. coli DH10B T1 resistant cells"
/clone_lib="Aspergillus flavus Normalized cDNA Expression Library"
/note="Vector: pBlueScript (SK-) (Stratagene), antibiotic selection marker: Carbenicillin; Site_1: Noti, at the 5

Db 792 ACCTCAGAGAGGCTAAGCGCTGAARAACTCTCCTGTTCTTTTGAACCTAGTAGAAGG 733

Qy 823 TTTTCTTTTGAACCCGATAGAGCGCGGTGAGATTAATAATGCGATGTAACACCCCGGC 882

Db 732 TTTCTTTTGGAGCCGACGAAA---TGGGATCCTAAAGTTTTCACAGAGTTTCCAGGA 676

Qy 883 TACACAAATATGTTCCAGAGTGCAGAGCGGACGATGATGAGCATTCGTTTCGAAAAC 942

Db 675 TACATAACCAAGACGGCATGAAATATTTTCGAGC-----CCCTTTATCGAAAC 625

Qy 943 CAGATTCCTAAAGAACGCGGCTTCGGGCCCTGCTGAAAGAGACAAATGCCCCAG 1002

Db 624 CAGATCCCGTGGAGCGGAAAGCAATGAGAGGTTTTCGAGCAGGTATTTCCCAA 565

Qy 1003 CTGGCAGACCGTCATTCAGCTTCGACGATTTGCTGGTGTGCGATACCGGAATCGC 1062

Db 564 ATCAGCAGCGGCTCCCTTGTATGCTGCCAAAATTTGCTGGTGCACGACACACCGATCGT 505

Qy 1063 GAATTCCTGATAGATGCACATCCGAGTACACAGTCTCTGTTGTTGGCTGTGGTGCAGC 1122

Db 504 CATTTCTTGATTTGTGAGCATCTGACATCAAGGGCTTGTATTTGGCAACCGGAGATTCT 445

Qy 1123 GGAAGAGGGTTAAATATCTGCCCTTCTATTGGGAATCTCATTTGTTGACGGAT----- 1175

Db 444 GGACAGGATTCAGTATATGCCATTAATTTGGCCACTAGATATCCGACCTCATTTGAAG 385

Qy 1176 --GGAAGGTAAAGTCCGCAAAAATTCACGAATTAATCAAGTGGAAACCCGCAATTGCG 1233

Db 384 GGAATATCGGATTTACAGAAACTTATCGCAAGCATGCGCTGGAGACAGAGATTGCC 325

Qy 1234 GCGAACCGTAACCTGGCGTGATCTCTCGGGGTTTGGCGGTCCAAATCGTGTGATGAT 1293

Db 324 GAGGGAAGAGATATATACACCTTCAAGGGAGGTTTGGTGGCTCAAAACAGAAATCAAGGAC 265

Qy 1294 TTTTCATGATGTAAGGAATGACCAATGTTC 1325

Db 264 CTCAATGAGCTAAGCGATTGACAAATAGCA 233

RESULT 3

CNS07AGE 906 bp DNA linear GSS 08-JUL-2001

LOCUS T3 end of clone BC0AA003H11 of library BC0AA from strain CBS 767 of

DEFINITION Debaromyces hansenii, genomic survey sequence.

ACCESSION AL436548

VERSION AL436548.1 GI:12219961

KEYWORDS GSS.

SOURCE Debaromyces hansenii (anamorph: Candida famata)

ORGANISM Debaromyces hansenii

REFERENCE 1 (bases 1 to 906)

AUTHORS Souciet J.L., Aigle M., Artiguenave F., Blandin G., Bolotin-Fukuhara M., Bon E., Bröttier P., Casaregola S., de-Montigny J., Dujon B., Durrens P., Lepingle A., Llorente B., Malpertuy A., Neuveglise C., Ozier-Kalogeropoulos O., Potier S., Saurin W., Tekala F., Toffano-Nicche C., Wesolowski-Louvel M., Wincker P. and Weissenbach J.

Genomic exploration of the hemiascomycetous yeasts: 1. A set of

yeast species for molecular evolution studies

JOURNAL FEBS Lett. 487 (1), 3-12 (2000)

MEDLINE 20584711

PUBMED 11152876

REFERENCE 2 (bases 1 to 906)

AUTHORS Lepingle A., Casaregola S., Neuveglise C., Bon E., Nguyen H., Artiguenave F., Wincker P. and Gaillardin C.

Genomic exploration of the hemiascomycetous yeasts: 14.

JOURNAL Debaromyces hansenii var. hansenii

FEBS Lett. 487 (1), 82-86 (2000)

MEDLINE 20584724

PUBMED 11152889

REFERENCE 3 (bases 1 to 906)

AUTHORS Genoscope.

TITLE
JOURNAL

Direct Submission
Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail :
segrif@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaromyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.

FEATURES

source

Location/Qualifiers
1. .906
/organism="Debaromyces hansenii"
/mol_type="genomic DNA"
/strain="CBS 767"
/variety="hansenii"
/db_xref="taxon:4959"
/clone="BC0AA003H11"
/clone_lib="BC0AA"
/notes="end : T3"

misc_feature

<15..>716
/notes="similar to 043029 [PUTATIVE FRUCTOSYL AMINO ACID
OXIDASE]"
/evidence="not_experimental"

ORIGIN

Query Match 7.0%; Score 99.2; DB 9; Length 906;
Best Local Similarity 50.8%; Pred. No. 3.3e-19;
Matches 335; Conservative 0; Mismatches 298; Indels 27; Gaps 3;
Qy 498 TGGCTGGGCACATCGAAGGAATGCTTAGTGGCAGCAGCAGCGAAGCAGCGCATGGG 557
Db 14 TGGGTGGGCATTTGCTAGGTGGCATTAGAGAAGCTGTTGAAGAATGCCGTAATAGG 73
Qy 558 TGTAAATTTGTTACTGGCACCCGCGAGGTCGTGTAGTCACTTAATCTTTGAAATAAA 617
Db 74 GGCTAAATTTGTTATTGATTCTGCTGAAGAACTCTCTCTCA-----GTGAAGATGG 124
Qy 618 CGATGTAAAGGTGCCGTGTAGGGCGGATGGCAAAATTTGGAGAGCGGNAACGTACATTCCT 677
Db 125 AGCATGTGTGGTGTTCACACTTCGAATGGAACATTTATTGAAGCTGATGAATATTAT 184
Qy 678 GTGTCTGGGGTAGCGCGGTGCTAGTCTCCTAGATTTCAAGAATCAACTTCGACCAACCGC 737
Db 185 TTGTGCTGGTGCAAAATTCATTCAGTTTTTTGAACCTTTGAGCAGCACTTTTGGCTAATG 244
Qy 738 TTGGACCCCTGGTACACATTCGCTTAAACCGGAAGAACGTGCGTTGTACAAAAATATACC 797
Db 245 CTATACCTTTGGACATATCAAACTCACTGATGATGAAGCTGCTTTGTTGAAGGGATGCC 304
Qy 798 GGTATCTTTAAACATCGAAGCGGGGTTTCTTTGAAACCGCATGAGGAGCGCGGTGAGAT 857
Db 305 AGTTGTTTAAATCTAGTAGCGGGTTTGTATTGAAACCGGACTTAA---CAATGAGAT 361
Qy 858 TAAATATGCGATCAACACCGCGGTCTACACAAATATGTCCTCAGAGTCGACGCGCAGAT 917
Db 362 TAAATCTGTAATGAGTTTCCAGGTATGTCAACATCGTAATGGAATTCAG----- 414
Qy 918 GATGAGCATTCCTGTTGAAAAAACCAGATTTCCAAAAAGAGCCGAAAGCGCGTTCGGGC 977
Db 415 -----TGCCTTTATTAAAGGATTCAATCCCAAGGAGGCTGAAGATCAATGAGAGC 466
Qy 978 CTGTCTGAAAGAGACAAATGCCAGCTGGCAGACCGCTCCATTTCAGCTTCGACCGATTG 1037
Db 467 ATTTTAAAGACAAGTTTTCCTCCCGAAATTTGCTGAAAGGGAGTTTTCATTAGCTAGAA 526
Qy 1038 CTGTGTGCGGATACCGCGAATCGCAATTCCTGTATAGATCGACATCGGAGTACCCAG 1097
Db 527 CTGTGACAGACACTCCAGATCGTCAATTTTGTGAGCATCCGGGGCATAAAA 586

```

QY      1098  TCTTGTTGGGTGCTGCGAGCGAAGAGGGTTTAAATATCTCGCTTCTATTGGGAA 1157
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      587  CTTGGTATTGGGTACTGCTGCTCAAGGTTTAAATACATGCCAAATGTTGAAA 646

RESULT 4
AA787493
LOCUS      n5a03al.r1 Aspergillus nidulans 440 bp mRNA linear EST 31-JUL-1998
DEFINITION vegetative cDNA lambda zap library Emericella nidulans cDNA clone
             n5a03al, mRNA sequence.
ACCESSION  AA787493
VERSION     AA787493.1 GI:2847724
KEYWORDS   Emericella nidulans (anamorph: Aspergillus nidulans)
SOURCE     Emericella nidulans
ORGANISM   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
             Eurotiales; Trichocomaceae; Emericella.
REFERENCE  1 (bases 1 to 440)
AUTHORS    Kupfer,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R.,
             Prade,R. and Roe,B.
TITLE      An Aspergillus nidulans EST Database
JOURNAL    Unpublished (1998)
COMMENT    Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
             Department of Chemistry and Biochemistry
             Advanced Center for Genome Technology, University of Oklahoma
             620 Parrington Oval, Norman, OK 73019, USA
             Tel: 405 325 4912
             Fax: 405 325 7762
             Email: broe@ou.edu
             We anticipate the future release of the cDNA clones to the Fungal
             Genetics Stock Center
             High quality sequence stop: 406.
FEATURES   source
             Location/Qualifiers
               1..440
                 /organism="Emericella nidulans"
                 /mol_type="mRNA"
                 /strain="FGSC A26"
                 /db_xref="taxon:162425"
                 /clone="n5a03al"
                 /tissue="vegetative mycelia, asexual structures"
                 /clone_lib="Aspergillus nidulans 24hr asexual
                 development and vegetative cDNA lambda zap library"
                 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
                 XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
                 3' end of cDNA cloned into XhoI site of pBluescript"

ORIGIN
Query Match      5.9%; Score 83.4; DB 1; Length 440;
Best Local Similarity 52.6%; Pred. No. 2.5e-14;
Matches 230; Conservative 0; Mismatches 201; Indels 6; Gaps 2;

QY      41  TCATAAGTCATCATCTCTCTGATCGTTGGTCGCGGACTTGGGCACTTCAACGGCTC 100
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      7  TCACCAAGACTCATCTATTCTCATCTATCGTGGTGGGTACTTGGGCTGCAGCACCGCTC 66

QY      101  TGCACCTCGCGCGCGGGGATACCAACGTTACCGTGTGTCGACCCCTATCTGTCCCTTA 160
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      67  TCCAGCTTGGCGGCGCGGCTACAAAACGTCACCGTCTTGAGCGCAGTCCGATACCGT 126

QY      161  GCGCCATCTCCGCGGAACGAGTGAACAAGCTATTAGCAGTGCCTATATTCCGAATA 220
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      127  CACCTCTGTGTCGCGGAATGATGTAACAAATCATGAGGAGGGCGGCAACCGTCCGACA 186

QY      221  ---ACAAAGACGAATCGAAGTGAATCAGATCTTTGGCGGAAGAGCGCTTTAAACGTTGGA 277
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      187  CAGACACCCCGAGCAATAGTCTGGNACCGGATGCACCATCATCCGCAATGCCTGGA 246

QY      278  AGAACGACCGGCTTTTCAAAACCGTATTATCATGATACGGGCCCTGTGATGCTGTGCT 337
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      247  AGACCGACCCCGGTATTTCGCCCTTCTTACCACCAACAGGTTTATTATGGCCGCGAGCC 306

```

```

QY      338  CGCAGAGCGGCTGGATCCCTCGGCGTCCGGGTACGTCCGGCGGAGGATCTTAATCTGG 397
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      307  AGGACGATGCAATGCCTCATGTTGACAAGTACATATTAGCACCTGCAGAGA---TAAACTCC 363

QY      398  TGGAACTTACCGCCCGGAGCAATTTCGTAACACTGCCCGGNAAGCGGTGTTCCAGAGTG 457
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      364  GGTACTCAACTCGGCGCGGATTTTCAGAGCTACAAATGCCGTGGCGGCTCTGACAGGGG 423

QY      458  ATTTTCGGGTTGGAAA 474
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      424  ACTTTTCGGCTGAAA 440

RESULT 5
BI187569/c
LOCUS      BI187569
DEFINITION a4902fs.fl Fusarium sporotrichioides Tri 10 overexpressed cDNA
             library Fusarium sporotrichioides cDNA clone a4902fs 3', mRNA
             sequence.
ACCESSION  BI187569
VERSION     BI187569.1 GI:14661248
KEYWORDS   Fusarium sporotrichioides
SOURCE     Fusarium sporotrichioides
ORGANISM   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
             Hypocreomycetidae; Hypocreales; mitosporic Hypocreales; Fusarium.
REFERENCE  1 (bases 1 to 397)
AUTHORS    Ren,Q., Tag,A., Peplow,A., Lai,H., Kupfer,C., Peterson,A.,
             Beremand,M. and Roe,B.
TITLE      Analysis of a Fusarium sporotrichioides EST database
JOURNAL    Unpublished (2001)
COMMENT    Other ESTs: a4902fs.r1
             Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
             Department of Chemistry and Biochemistry
             Advanced Center for Genome Technology, University of Oklahoma
             620 Parrington Oval, Norman, OK 73019, USA
             Tel: 405 325 4912
             Fax: 405 325 7762
             Email: broe@ou.edu
             Contact Dr. Marian Beremand regarding clone availability included
             is the best homolog from a blastx search of Genbank nr 04-09-01
             234 3e-19 gi|2661130|gb|AAB882 (AF035700) fructosyl
             amine:oxygenoxidoredo
             Seq primer: M13-20
             High quality sequence stop: 253.
FEATURES   source
             Location/Qualifiers
               1..397
                 /organism="Fusarium sporotrichioides"
                 /mol_type="mRNA"
                 /strain="Tri 10"
                 /db_xref="taxon:5514"
                 /clone="a4902fs"
                 /clone_lib="Fusarium sporotrichioides Tri 10 overexpressed
                 cDNA library"
                 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
                 XhoI; 5' end of cDNA cloned into EcoRI site of
                 pBluescript; 3' end of cDNA cloned into XhoI site of
                 pBluescript"

ORIGIN
Query Match      5.1%; Score 72.6; DB 4; Length 397;
Best Local Similarity 53.3%; Pred. No. 6.2e-11;
Matches 153; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY      1011  CGTCCATTAGCTTCGCACGCAATTTGCTGGTGTCCGATACCGCGAATCGCAATTCCT 1070
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      397  CAGACCATTGTCACGCTCGAATCTGTTGGTGTGTGTATACATATGACAGAAACTTCCT 338

QY      1071  GATAGATGCATCCCGCAGTACCACAGTCTTGTGTTGGGCTGTGTCCGAGCGGAGAGG 1130
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      337  CATTACATATCATCCACGACATCCATCGCTCGTTCATCGCTTCGGGTGACTGCGGACAGG 278

QY      1131  GTTTAAATATATGCGCTTCTATTGGGAATCTCATTTGTGACGCGATGGAAGTAAAGTGCC 1190
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db      277  ATCAAGCATTACATCTATTCGCAATTCATCTCGATTGTATGAGGTAGATTGGA 218
Qy      1191  GCAAAATTCAGAAATTAATCAAGTGAACCCGGAACATTCGGCGGAACCGTAACTGCGG 1250
Db      217  AGAGAGATTTCGCAAGTTTGGACATGAGACCGGAGAGTTTACTGAGTTTGGGGTAA 158
Qy      1251  TGATACTCTGGGGGTTTTCGGCGTCCAAATCGTGTGATGATTTTC 1297
Db      157  CGATCCTTAGATGATTTGGTGTGAGTACTATCATGATCTGC 111

RESULT 6
CN241854
LOCUS      677 bp      mRNA      linear      EST 09-APR-2004
DEFINITION      EST007722 Mycelium and yeast cells from Paracoccidioides
                  brasiliensis Paracoccidioides brasiliensis cDNA, mRNA sequence.
ACCESSION      CN241854
VERSION        CN241854.1 GI:46345598
KEYWORDS
SOURCE
ORGANISM      Paracoccidioides brasiliensis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; mitosporic Onygenales; Paracoccidioides.
REFERENCE      1 (bases 1 to 677)
AUTHORS        Felipe, M.S.S., Carvalho, M.J.A., Andrade, R.V., Arraes, F.B.M.,
                  Simoes, I.C., Andrade, E.V., Maranhao, A.Q., Torres, F.A.G.,
                  Jesuino, R.S.A., Kyaw, C.M., Moraes, L.M.P., Nicola, A., Pereira, M.,
                  Silva-Pereira, I., Anjos, D.A.S., Sandes, E.F.O., Inoue, M.K.,
                  Walter, M.E.M.T., Soares, C.M.A. and Brigo, M.M.
TITLE          Metabolic features of Paracoccidioides brasiliensis cell
                  differentiation as assessed by transcriptome analysis
JOURNAL
COMMENT        Unpublished (2004)
Contact: Felipe MSS
Laboratory of Molecular Biology
Institute of Biology - University of Brasilia
Campus Universitario, Asa Norte, Brasilia, DF 70910-900, BRA
Tel: 55 61 307 2423
Fax: 55 61 349 8411
Email: mauei@unb.br
Seq primer: T7 Sequencing primer.

FEATURES
source
1..677
/organism="Paracoccidioides brasiliensis"
/mol_type="mRNA"
/strain="Pb01"
/db_xref="taxon:121759"
/clone_lib="Mycelium and yeast cells from Paracoccidioides
brasiliensis"
/notes="Pb Lambda Zap Express Library"

ORIGIN
Query Match      4.7%; Score 66.2; DB 7; Length 677;
Best Local Similarity 51.2%; Pred. No. 7.7e-09;
Matches 155; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

Qy      933  CGAAAAAACCCAGATTCCAAAAGAACCGGCGTTTCGGCGCTCTGTGAAGAGAC 992
Db      243  CGGAGATGTCCGATCCGATCCTAGGGGAGGAGGATGTAGAGCCGCTGAGAGAGAT 302
Qy      993  AATGCCCCAGCTGGCAGACCGTCCATTACGTTTCGACGCGATTCGTGGTGGCGATAC 1052
Db      303  GCTGCCCGAGCTTTGGCAGAGCGGGCTTTGTGAACACCGGAGTGTGTGGTATCCGGATAC 362
Qy      1053  CGCGAATCGGGAATTCCTGTATAGATCGACATCGCAGTACACAGTCTGTGTGGGCTG 1112
Db      363  ACCACAGGGGAGCTTTCTAGTAACTACACACCACTACCGGGGCTCTCTCTCGCAAC 422
Qy      1113  TGGTGGCAGCGGAAGAGGGTTTAAATATCTGCCTTCTATTGGGAATCTCAATTGTGACGC 1172
Db      423  CGGTGGCAGCGGGCAGCGGGTTTAAATTCCTCCCGTCTCGGGGAGAAGATCGTCGATGC 482
Qy      1173  GATGGAAGGTAAGTGGCCGCAAAAATTCAGGAATTAATCAAGTGGGAACCGGACATGC 1232

```

```

Db      483  GATACAGGACCGTTTGATCTGGAGTTGCAGAGTATATGGCGTGGACTAGTCTACTTC 542
Qy      1233  GGC 1235
Db      543  GTC 545

RESULT 7
BI187570
LOCUS      414 bp      mRNA      linear      EST 10-JUL-2001
DEFINITION      a4g02fs.r1 Fusarium sporotrichioides Tri 10 overexpressed cDNA
                  library Fusarium sporotrichioides cDNA clone a4g02fs 5', mRNA
                  sequence.
ACCESSION      BI187570
VERSION        BI187570.1 GI:14661249
KEYWORDS
SOURCE
ORGANISM      Fusarium sporotrichioides
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; mitosporic Hypocreales; Fusarium.
REFERENCE      1 (bases 1 to 414)
AUTHORS        Ren, Q., Tag, A., Peplow, A., Lai, H., Kupfer, C., Peterson, A.,
                  Beremand, M. and Roe, B.
TITLE          Analysis of a Fusarium sporotrichioides EST database
JOURNAL
COMMENT        Other_ESTs: a4g02fs.fl
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
Contact Dr. Marian Beremand regarding clone availability Included
is the best homolog from a blastx search of Genbank nr 04-09-01
217 3e-29 gi|2661130|gb|AAB882 (AF035700) fructosyl
amine:oxygenoxidoredu
Seq primer: T3.

FEATURES
source
1..414
/organism="Fusarium sporotrichioides"
/mol_type="mRNA"
/strain="Tri 10"
/db_xref="taxon:5514"
/clone="a4g02fs"
/clone_lib="Fusarium sporotrichioides Tri 10 overexpressed
cDNA library"
/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; 5' end of cDNA cloned into EcoRI site of
pBluescript; 3' end of cDNA cloned into XhoI site of
pBluescript"

ORIGIN
Query Match      4.6%; Score 65.2; DB 4; Length 414;
Best Local Similarity 54.7%; Pred. No. 1.4e-08;
Matches 176; Conservative 0; Mismatches 138; Indels 8; Gaps 2;

Qy      31  GCTCTCGCCGCTCACTAAGTCATCTCTCTGATCGTTGGTCCGCGACTTGGGGCACC 90
Db      12  GCTTCATCTCTTACCAAGAGTCCAAATCTCATCGTTGGAGGTGGGACATGGGATGT 71
Qy      91  TCAACGGCTCTGCACCTTCGCGCGCGGATATACCAACGTTTACCGTGTGGACCCCTAT 150
Db      72  TCCACAGCTCTTCATCTAGCTCGTGGAGGCTATACCAAGTCACACTTCTCGA--TGTTA 129
Qy      151  CTGTCTCCTAGCCCATCTCCGCGCGGAAAACGAGTGAACAAAGTCATTAG-----CAGT 204
Db      130  ACCATCCCTTCCACCAATCTCTCGCGGCCATGATGTCAACAAGATTGCAAGTGGATTGAAT 189
Qy      205  GGCCAATATTTCGAATCAACAAGACGAATCGAAGTGAATGAGATCTTGGCGGAAGAGCG 264
Db      190  GTCCCGATTAATTAAGAGACGATGAAGATGCAATCTGGCAATCCCTCACCCTACGCCCA 249

```

```
QY 265 TTAAACGGTTGGAAGAACGACCGCTTTTCAACCGTATTATCATGATACGGCCCTGCTG 324
Db 250 GCTCAGGGCTGGCTACATGATCCAGTCTTCAAAACCATATCTACCAACGACAGGCTATGTC 309

QY 325 ATGCTCTGCTTCTCGCAGGAGG 346
Db 310 ATCTCTGCTTCAACGCGCAAAG 331

RESULT 8
AJ636269
LOCUS AJ636269
DEFINITION AJ636269 Mga Mycosphaerella graminicola cDNA clone mgal1757f, mRNA
sequence.
ACCESSION AJ636269
VERSION AJ636269.1
KEYWORDS GI:47029326
SOURCE EST.
ORGANISM Mycosphaerella graminicola
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et
Chaetothryiomycetes incertae sedis; Mycosphaerellaceae;
Mycosphaerella.
REFERENCE 1 (bases 1 to 599)
AUTHORS Keon, J. P. R., Bailey, A. M. and Hargreaves, J. A.
TITLE A group of expressed cDNA sequences from the wheat fungal leaf
blotch pathogen, Mycosphaerella graminicola (Septoria tritici)
JOURNAL Fungal Genet. Biol. 29 (2), 118-133 (2000)
MEDLINE 20374020
PUBMED 10919380
COMMENT Contact: Keon J
Plant Pathogen Interactions Division,
Rothamsted Research,
Harpenden, Herts, UNITED KINGDOM
Tel: +44(0)1582 763133
Fax: +44(0)1582 760981
Email: john.keon@bbsrc.ac.uk
Insert Length: 800 Std Error: 100.00
Seq primer: M13 reverse.

FEATURES
source
1..599
/organism="Mycosphaerella graminicola"
/mol_type="mRNA"
/strain="Strit"
/db_xref="taxon:54734"
/clone="mgal1757f"
/clone_lib="Mga"
/notes="Vector: pSPORT1; Library constructed from cultures
grown in Czepek Dox for 3 days then Glucose Yeast Nitrogen
Base for 1 day, shift of nitrogen source from nitrate to
ammonium ions"

ORIGIN
Query Match 4.6%; Score 64.6; DB 1; Length 599;
Best Local Similarity 53.8%; Pred. No. 2.4e-08;
Matches 133; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 937 AAACCCAGATTCCAAAGACGCGAAACCGCGTTTCGGCCCTGCTGAAGACAAATG 996
Db 217 AACATGCAATCCCAACGGAAGAGAGCGCTCGCTCGCGTCTTCTGCGAAAAGCCATA 276
QY 997 CCCAGCTGGCAGACCGTCCATTCACTTCGACGCAATTTGCTGGTGTGCCGATACCGCG 1056
Db 277 CCTCGATGGCGACCGACCATCGGCCCAACACGGGATATGCTGGTACACGGACCGTC 336

QY 1057 AATCGCGAATCTCTGATAGATCGACATCCGAGTACACAGTCTTGTTGGGCTGTGGT 1116
Db 337 TCTGGAGACTGGTTGATCATCAGTACCATCCGAAGTATTCGGGACTCTTCGTGCGCACGGC 396
QY 1117 GCAGCGGAAGAGGGTTTAATATCTGCTTCTATTGGGATCTCATTTGTTGAGCGATG 1176
Db 397 GGCTCGGGACATGGATATAAATCTTTCGCCGATCATTTGGAGGGAGAAATTTGGATGTGATT 456
```

```
QY 1177 GAAGGTA 1183
Db 457 GAGGGCA 463

RESULT 9
BM869701
LOCUS BM869701
DEFINITION BM869701.2 GI:30405219
ACCESSION BM869701
VERSION BM869701.2
KEYWORDS GI:30405219
SOURCE EST.
ORGANISM Magnaporthe grisea (anamorph: Pyricularia grisea)
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
REFERENCE 1 (bases 1 to 628)
AUTHORS Ebbole, D. J., Yuan, J., Thomas, T. L., Bobrowicz, P., Lu, G.,
Bhatterai, K., and Dean, R. A.
TITLE Expressed sequence tags from the rice blast fungus, Magnaporthe
grisea
JOURNAL Unpublished (2002)
COMMENT On Mar 7, 2002 this sequence version replaced gi:192377383.
Contact: Ebbole DJ
Department of Plant Pathology & Microbiology
Texas A&M University
Peterson Bldg. MS2132, College Station, TX 77843-2132, USA
Tel: 979 845 4831
Fax: 979 845 6483
Email: d-ebbole@tamu.edu
Chromatogram file of this sequence is available, see contact
person; Best nr hit (April. 22, 2003) gb|EAA30984.1| hypothetical
protein [Neurospora crassa] 194 6e-49
PCR Primers
FORWARD: T3 primer
BACKWARD: T7 primer
Plate: mgns007 row: F column: 15
Seq primer: T3.

FEATURES
source
1..628
/organism="Magnaporthe grisea"
/mol_type="mRNA"
/strain="Guy11"
/db_xref="taxon:148305"
/clone="mgns007xF15"
/sex="Mat1-2 hermaphrodite"
/cell_type="mycelium"
/clone_lib="Magnaporthe grisea NS Uni-Zap XR Library"
/notes="Vector: pBluescriptSK-; Site 1: EcoRI; Site 2:
XhoI; Unidirectional cloning. EcoRI side has T3 primer and
predominantly 5' reads. T7 primer on XhoI side of insert.
Nitrogen starvation library. Cells were inoculated into
minimal medium and grown for two days with shaking (150
rpm) at room temperature. Culture was harvested, blended,
inoculated into minimal medium as above for 24 h. Cells
were harvested, washed with water and inoculated into
minimal medium base lacking nitrogen source for 6 h.
Sequences were processed by one of two methods. Where a
full-length alignment to the M. grisea genome sequence was
available, the EST sequence was trimmed according to the
alignment, otherwise sequence quality was assessed using
phredPhrap version 991019 and trimmed according to phd
files (0.05) and for vector seqs."

ORIGIN
Query Match 4.6%; Score 64.6; DB 4; Length 628;
Best Local Similarity 56.3%; Pred. No. 2.4e-08;
Matches 121; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 992 CAATGCCCGAGCTGGCAGACCGCTCCATTGAGTTCGACCGCATTTGCTGTGCGCGATA 1051
Db 217 CGATCTCTCGATTTCATGACCGCGCCCTTCACTTCAACCCCGCTATGCTGTACACCGACA 276
```

QY	1052	CCGCGAATCCGCGAATTCCTGATGATCGACATCCGCGAGTACCAAGTCTCTGTGTGTGGGCT	1111
Db	277	CAAAAGTCGGCGCATTTTCATCATCGATTACCATCTCGAGTGGGAAGGTCTTTTCTTCGCCCA	336
QY	1112	GTGGTGGACGCGGAAGAGGGTTTAAATATCTGCCTTCTATTGGGAATCTCATTTGTTGACG	1171
Db	337	CTGGTGTAGCGCGCATGGTTTCAAGTCTTTCGGGTTCTTTGGAGAACGGATTGCCGACT	396
QY	1172	CGATGGAAGTAAAGTGGCCGCAAAAATTCACGAA	1206
Db	397	GTATCGCATATGATCGACCGTCAGAATTTCGGGAA	431

RESULT 10	
CF188389/c	
LOCUS	459 bp mRNA linear EST 08-AUG-2003
DEFINITION	k2d07j2.f1 Cryptococcus neoformans strain B3501 Cryptococcus neoformans var. neoformans cDNA clone k2d07j2.3, mRNA sequence.
ACCESSION	CF188389
VERSION	CF188389.1 GI:33510258
KEYWORDS	EST.
SOURCE	Cryptococcus neoformans var. neoformans (Filobasidiella neoformans var. neoformans)
ORGANISM	Cryptococcus neoformans var. neoformans Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
REFERENCE	1 (bases 1 to 459)
AUTHORS	Kupfer,D.M., Drabenstot,S.D., Buchanan,K.L., Lai,H., Dyer,D.W., Roe,B.A. and Murphy,J.W.
TITLE	Comparison of highly conserved intronic and exonic elements associated with splicing among five diverse fungal organisms Unpublished (2003)
JOURNAL	Other ESTs: k2d07j2.r1
COMMENT	Contact: Murphy, JW Department of Microbiology and Immunology University of Oklahoma Health Sciences Center Oklahoma City, OK 73190, USA Tel: 405-271-2133 ex2133 Email: juneann-murphy@ouhsc.edu This clone is available from the Fungal Genetics Stock Center, Phone 913-588-7044, http://www.fgsc.net/ Contact Dr. Bruce Roe (brose@ou.edu , www.genome.ou.edu) for sequencing questions Contact Dr. Juneann Murphy(juneann-murphy@ouhsc.edu) for library information.

Qy	1065	ATTCTGATAGATCGACATCCGAGTACCAACAGTCTTGTGTGGCTGTGTGTCGAGCGG	1124
Db	321	CTGGCTACTATCGAGACATCCCAAGTGGAAAGTGTGGTGTGGCTACGCGAGATATGTG	262
Qy	1125	AAGAGGGTTTAAATATCTGCTTCTATATCGGAATCTCATTTGTGACGCGATGGAAGGTAA	1184
Db	261	GCACACTTTTAAATGCTGCCTGTGGTGGCGCCAAAGTTGCTGACCTCATCGAAGCAA	202
Qy	1185	AGTCCGCGAAAAATTCAGCAATTAATCAAGTGGAAACCCGG	1225
Db	201	AATGTCGAGCAGAGAAAGTATGCTTTGGAGGTGGAGACCTG	161

RESULT	11
BG465351	
LOCUS	EM1_75_C06_g1_A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA
DEFINITION	sequence.
ACCESSION	BG465351
VERSION	BG465351.1
KEYWORDS	EST.
SOURCE	G1:13394327
ORGANISM	Sorghum bicolor (sorghum)
	Sorghum bicolor
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Sorghum. 1 (bases 1 to 422) Reid,S.P., Cordonnier-Pratt,M.-M., Gingle,A. and Pratt,L.H. An EST database from Sorghum: developing embryos Unpublished (2000) Contact : Cordonnier-Pratt MM JOURNAL COMMENT

FEATURES	Location/Qualifiers
source	1..422
	/organism="Sorghum bicolor"
	/mol_type="mRNA"
	/db_xref="taxon:4558"
	/clone_lib="Embryo 1 (EM1)"
	/notes="Organ: Embryos germinated for 24 hr; Vector: pBluescript II from Lambda Zap II; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A_RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."


```

QY 1124 GAAGAGGGTTTAATATCTGCCTTATTGGGAATCTCAATGTTGACGGATGGAAGGTA 1183
Db 200 GGCATGGCTACAGTTCGTGCCGAAGATTGGCGAGAGGATCGTGGATGTAATGAAGCGCC 259

QY 1184 AAGTCCGCGAATAATTCAGGAATTAATCAAGTGGACCGGACATTTGGCGGAACCGTA 1243
Db 260 AACCACGGGATGACTTGGGTACTGAAATTCAGGAAGAAGTGGTCATGCGCGAAGCAGAAT 319

QY 1244 ACTGGCGTGATCTCTGGGCGGTTTTGGCGGTCCAAATCGTGATGGATTTTCATGATG 1303
Db 320 ACCACTTTGATCACTCTGCAACGACTGGAGAGTGACGAAGGATGATGATTGTTG 379

QY 1304 TGAAGGAATGGACCAATGTTCACTA 1328
Db 380 AGGAAGACGAGGAGGCGGAGAA 404

RESULT 12
CF871025 758 bp mRNA linear EST 31-OCT-2003
LOCUS
DEFINITION
Hypocrea jecorina cDNA clone tric025xol7, mRNA sequence.
ACCESSION
CF871025
VERSION
CF871025.1 GI:38125707
KEYWORDS
EST.
SOURCE
Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM
Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
REFERENCE
Diener, S.E., Dunn-Coleman, N., Foreman, P., Houfek, T.D.,
Teunissen, P.J.M., van Solingen, P., Dankmeyer, L., Mitchell, T.K.,
Ward, M. and Dean, R.A.
Characterization of the protein processing and secretion pathways
in a comprehensive set of expressed sequence tags from Trichoderma
reesei.
JOURNAL
FEMS Microbiol. Lett. 230 (2), 275-282 (2004)
COMMENT
Contact: Ralph A. Dean
Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
Fax: 919-513-0024
Email: ralph.dean@ncsu.edu
Seq primer: LT-F1 primer.
FEATURES
source
1..758
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clone="tric025xol7"
/dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 6 October
2003"
/notes="Vector: pREP3Y; Site 1: Not 1/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."
ORIGIN
Query Match 3.9%; Score 55.2; DB 7; Length 758;
Best Local Similarity 60.8%; Pred. No. 2.3e-05;
Matches 90; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 53 CATCTCTCGATCGTTGGTGGCGGACCTTGGGGACCTCAACGGCTCTGCACCTGCGGC 112
Db 346 CGTCAGTTCTGATTGTCGGTGTGGTAACTTCGGCGCGCAACGGCTTTGAGTCTGGCCA 405

QY 113 GCCCGGATATACCAAGTTACCGTGTGGACCCCTATCTGTCCTTAGCGCCATCTCCG 172
Db 406 GAAGGGGAGGGTTCAAGGTCAACATCATTTGACACGGCCCCCTTCCCCAACCCAGGCGCG 465

QY 173 CCGGAACGACGCTGAACAAAGTCATTAG 200

```

```

Db 466 CCTCCACGACATCAACRAGATCGTTCC 493

RESULT 13
CB901239 816 bp mRNA linear EST 02-JUL-2003
LOCUS
DEFINITION
jecorina cDNA clone tric025xol7, mRNA sequence.
ACCESSION
CB901239
VERSION
CB901239.1 GI:30115897
KEYWORDS
EST.
SOURCE
Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM
Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
REFERENCE
Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S.,
Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J.,
Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C.,
Olivares, H.A., Teunissen, P.J., Yao, J. and Ward, M.
Transcriptional regulation of biomass-degrading enzymes in the
filamentous fungus Trichoderma reesei
J. Biol. Chem. 278 (34), 31988-31997 (2003)
JOURNAL
MEDLINE
PUBMED
12788920
COMMENT
Contact: Pamela K. Foreman
Genencor Intl.
925 Page Mill Road, Palo Alto, CA 94304, USA
Tel: (650) 846-7635
Fax: (650) 621-7817
Email: Pforeman@genencor.com
Seq primer: LT-F1 primer.
FEATURES
source
1..816
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clone="tric025xol7"
/dev_stage="mycelia"
/clone_lib="T.reesei mycelial culture, Version 3 april"
/notes="Vector: pREP3Y; Site 1: Not 1/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."
ORIGIN
Query Match 3.9%; Score 55.2; DB 6; Length 816;
Best Local Similarity 60.8%; Pred. No. 2.4e-05;
Matches 90; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 53 CATCTCTCGATCGTTGGTGGCGGACCTTGGGGACCTCAACGGCTCTGCACCTGCGGC 112
Db 404 CGTCAGTTCTGATTGTCGGTGTGGTAACTTCGGCGCGCAACGGCTTTGAGTCTGGCCA 463

QY 113 GCCCGGATATACCAAGTTACCGTGTGGACCCCTATCTGTCCTTAGCGCCATCTCCG 172
Db 464 GAAGGGGAGGGTTCAAGGTCAACATCATTTGACACGGCCCCCTTCCCCAACCCAGGCGCG 523

QY 173 CCGGAACGACGCTGAACAAAGTCATTAG 200
Db 524 CTCTCCACGACATCAACRAGATCGTTCC 551

RESULT 14
CB901049 822 bp mRNA linear EST 02-JUL-2003
LOCUS
DEFINITION
jecorina cDNA clone tric025xe08, mRNA sequence.
ACCESSION
CB901049
VERSION
CB901049.1 GI:30115707
KEYWORDS
EST.
SOURCE
Hypocrea jecorina (anamorph: Trichoderma reesei)

```


ORGANISM	Hypocrea jecorina			
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.			
AUTHORS	1 (bases 1 to 822) Foreman,P.K., Brown,D.E., Dankmeyer,L., Dean,R., Diener,S., Dunn-Coleman,N.S., Goedegebuur,F., Houfek,T.D., England,G.J., Kelley,A.S., Meerman,H.J., Mitchell,T., Mitchinson,C., Olivares,H.A., Teunissen,P.J., Yao,J. and Ward,M.			
TITLE	Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus Trichoderma reesei			
JOURNAL	J. Biol. Chem. 278 (34), 31988-31997 (2003)			
DOI	22803314			
PUBMED	12788920			
COMMENT	Contact: Pamela K. Foreman Genencor Intl. 925 Page Mill Road, Palo Alto, CA 94304, USA Tel: (650) 846-7635 Fax: (650) 621-7817 Email: Pforeman@genencor.com Seq primer: LT-F1 primer. Location/Qualifiers 1..822 /organism="Hypocrea jecorina" /mol_type="mRNA" /strain="QM6a" /db_xref="taxon:51453" /clone="tric025xe08" /dev_stage="mycelia" /clone_lib="T.reesei mycelial culture, Version 3 april" /note="vector: pREP4Y; Site_1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."			
ORIGIN	Query Match 3.9%; Score 55.2; DB 6; Length 822; Best Local Similarity 60.8%; Pred. No.2.4e-05; Matches 90; Conservative 0; Mismatches 58; Indels 0; Gaps 0;			
Qy	53	CATCTCTCCGTATCGTTGGTCCGGGACCTCGGGCACCTCAACGGCTCTGCACCTCGGCGC	112	
Db	420	CGTCAGTTCTGATTGTCGGTGTCTGTAACCTTCGGCGCGCAACGGCTTTGAGTCTGGCCA	479	
Qy	113	GCOCGGGATATACCAAGTTACCGTGTGGACCCCTATCTGTCTCCCTAGGCCCATCTCCG	172	
Db	480	GAAGGGGAGGGTTCAAGTCAACCATCTTGACACGGCCCTCTCCCAACCCAGGGCGC	539	
Qy	173	CCGGAACGACGTCGAACAAAGTCATTAG	200	
Db	540	CTCTCCACGACATCAACAAAGATCGTTTCG	567	
RESULT 15				
CF870843				
LOCUS	CF870843 822 bp mRNA linear EST 31-OCT-2003			
DEFINITION	tric025xe08.b1 T.reesei mycelial culture, Version 6 October 2003			
ACCESSION	Hypocrea jecorina cDNA clone tric025xe08, mRNA sequence.			
VERSION	CF870843			
KEYWORDS	CF870843.1 GI:38125525			
SOURCE	EST.			
ORGANISM	Hypocrea jecorina (anamorph: Trichoderma reesei)			
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.			
AUTHORS	1 (bases 1 to 822) Diener,S.E., Dunn-Coleman,N., Foreman,P., Houfek,T.D., Teunissen,P.J.M., van Solingen,P., Dankmeyer,L., Mitchell,T.K., Ward,M. and Dean,R.A.			
TITLE	Characterization of the protein processing and secretion pathways in a comprehensive set of expressed sequence tags from Trichoderma reesei			
JOURNAL	FEMS Microbiol. Lett. 230 (2), 275-282 (2004)			
COMMENT	Contact: Ralph A. Dean Fungal Genomics Laboratory			

THIS PAGE BLANK (USPTO)